HERC's inpatient average cost datasets for VA Care

Version 4: Fiscal Years 1998-2002

Todd H. Wagner, PhD
Shuo Chen, PhD
Wei Yu, PhD
Paul G. Barnett, PhD

July 30, 2003

Health Economics Resource Center

Acknowledgments

A number of individuals have made this research possible and to them we owe our thanks. First, we would like to acknowledge our expert panel: Ann Hendricks, Denise Hynes, and Terri Menke. Their insight, experience and advice helped to produce a higher quality product. We would also like to thank Steve Wright for providing us with the Medicare data and John Rodgers for his SAS programs, many of which we still use. Lastly, we would like to acknowledge Health Services Research and Development Service (HSR&D) and the Cooperative Studies Program (CSP) for funding this research.

Table of Contents

Chapter 1. Overview . . . 1
1.1 Departures from past years' methods . . . 2
1.1.1 Acute medical-surgical short stay hospitalizations . . . 2
1.1.2 Categories of inpatient care . . . 2
1.1.3 Nursing home care . . . 2
1.2 Organization of paper . . . 3

Chapter 2. Cost Distribution Report . . . 4
2.2 CDR units and unit costs . . . 6
Chapter summary . . . 7

Chapter 3. VA Inpatient Databases . . . 8
3.1 VA Utilization Datasets . . . 8
3.1.2 PTF acute care bedsection discharge file (PTF bedsection) . . . 9
3.1.3 PTF Acute Census files . . . 9
3.1.4 PTF Extended care files . . . 10
3.1.5 Observation Bed files . . . 10
Chapter summary . . . 11

Chapter 4. Merger of cost and utilization databases . . . 12
4.1 Excluded facilities . . . 12
4.2 Facility mergers . . . 12
4.3 Definition of patient care unit . . . 13
4.4 Merger of cost and inpatient utilization data . . . 14
4.5 Data reconciliation . . . 15
4.6 Daily rate . . . 16
Chapter summary . . . 17

Chapter 5. The cost of non-medical/surgical inpatient care . . . 18
5.1 What is non-medical/surgical inpatient care? . . . 18
5.2 Cost methodology for non-medical/surgical care . . . 18
5.2.1 Leave and pass days . . . 18
5.2.2 Local outlier costs . . . 18
5.2.3 Why local rates at all? . . . 19
5.2.4 Adjusting for case-mix . . . 19
Chapter summary . . . 20

Chapter 6. The cost of nursing home care . . . 21
6.1 Case mix index . . . 21
6.2 Patient level case-mix . . . 21
6.2.1 RUG score at discharge . . . 22
6.2.2 Average WWU . . . 23
6.2.3 Exceptions . . . 26
6.3 Case-mix index of a medical center . . . 26

6.4 National case-mix index 26
6.5 Relative Value Unit (RVU) 26
6.5.1 Average case-mix adjusted local cost 27
6.5.2 Average case-mix-adjusted national cost 27
6.6 Distribution of case-mix 27
Chapter summary 28

Chapter 7. The cost of acute medical-surgical hospitalizations 29
7.1 Making an acute medical-surgical inpatient discharge database 29
7.2 Selecting the DRG and the relative value associated with a DRG 32
7.3 Length of stay 32
7.4 Building the cost function 33
7.4.1 Data 33
7.4.2 Cost adjusted charges 34
7.4.3 The dependent variable 35
7.4.4 Length of stay 35
7.4.5 Individual DRG intercepts or DRG weights 36
7.4.6 Final model 36
7.5 Observation days 38
7.6 Negative or implausible costs 39
7.7 Reconciling to the CDR 39
7.8 Stability of the cost function over time 41
Chapter summary 43

Chapter 8. User's Guide 44
8.1 Summary of methods 44
8.1.1 Categories of inpatient care 44
8.1.2 Acute medical-surgical care 44
8.1.3 Nursing home care 45
8.1.4 Non medical/surgical categories 45
8.2 Assumptions in the average cost dataset 45
8.2.1 Data used in the cost function 46
8.2.2 The cost of observation stays 46
8.2.3 Costs for high and low-cost procedures 46
8.2.4 Implicit trimming of outliers 46
8.2.5 Model estimates and negative costs 46
8.2.6 VISN administrative costs 47
8.3 Using the average cost dataset 47
8.3.1 Two important variables: source and flag 47
8.3.2 Discharge dataset 48
8.3.3 Acute medical-surgical dataset 49
8.3.4 Non medical-surgical dataset 49
8.4 When not to use the average cost dataset 50
8.4.1 Effects not detected in this cost estimate 50
8.4.2 Comparison of medical center efficiency 50
8.4.3 Point estimates versus variance estimates 51

8.5 Duplicates 52

References 53

Appendix A
Reconciliations for 1998 54

Appendix B
Flow diagram for inpatient care 58

Appendix C
VHA directive on observation beds 61

Appendix D
Aggregating the bedsection files 65

Tables

Table 2.1: Cost Distribution Accounts (CDAs) in the Cost Distribution Report Inpatient Services 5
Table 4.1: Excluded Facilities 12
Table 4.2: Facility Consolidations in 1997- 13
Table 4.3: Categories of Inpatient Care 15
Table 4.4: Median facility cost per day of stay for inpatient care 16
Table 6.1: RUG II classification and Wage-Weighted Work Units 22
Table 6.2: Distribution of RVUs at Patient and Institutional Levels in FY98 27
Table 7.1: Full model based on 50% random sample of Medicare data 37
Table 7.2: Correlations between estimated costs and actual costs for the full model and for three outlier restricted models 38
Table 7.3: Stability of regression coefficients with 1994, 1995 and 1996 Medpar data 41
Table 7.4: Pair wise Correlations in predicted costs compared to 1996 costs adjusted charges 42
Table 8.1: Included and excluded costs 45
Table 8.2 The three average cost datasets for FY98 47
Table 8.3 Using the three average cost datasets 50
Table 8.4: The cost function's effect on the variation of the estimated costs 51

Figures

Figure 3.1: VA inpatient data files 8
Figure 6.1: Number of possible assessments used to calculate
an average WWU in 8 situations 24
Figure 7.1: Accumulating contiguous acute medical-surgical bedsection stays 30
Figure 7.2: Distribution of cost adjusted charges 35
Figure 7.3: Difference between FY view and discharge view 40

Chapter 1. Overview

The U.S. Department of Veterans Affairs (VA) provided health care to veterans at 146 medical centers in 1998. Abstracts of all inpatient and outpatient utilization are centrally available at the Austin Automation Center. Tracked as part of these abstracts are the utilization data. However, no encounter-level charge or cost information is present. This is because the VA does not routinely generate patient bills. Consequently, VA researchers have not had economic data to estimate the cost of health care encounters.

In 1999, the VA funded the Health Economics Resource Center (HERC) to adapt existing cost methodologies (Barnett, Chen, & Wagner, 2000) and expand methods where possible and necessary. The current methodology, described in detail in this paper, is evolving and continues to improve over time. Input from users is crucial so that improvements can be made. We welcome all suggestions.

This report describes HERC's method for estimating the cost of VA health care encounters in fiscal years 1998-2001.[1] Our goal was to develop a set of long-term costs that could be used in cost-effectiveness analysis. By long-term we mean that all costs are variable.

Known as the "average cost" method, we assume that every health care encounter has the average cost of all encounters that share its same characteristics. While this assumption limits the accuracy of the cost estimates, especially for outliers, this is the only available method of generating a comprehensive set of encounter-level estimates of all patient care provided by VA. The average cost method relied on the following assumptions:

- To find the cost of rehabilitation, blind rehabilitation, spinal cord injury, psychiatric, substance abuse, intermediate medicine, domiciliary, and psychosocial residential rehabilitation stays, we found the average cost of a day of stay, and multiplied it by length of stay to estimate the cost of care. This makes the assumption that every day of stay has the same cost, that is, that costs are directly proportionate to the length of stay. This type of care is hereafter referred to as non-medical/surgical.

- To find the cost of acute medical-surgical hospital care, we built a cost function using relative value units (RVUs) from the non-VA sector. These RVUs were the Diagnosis Related Group (DRG) weights used by the Centers for Medicare and Medicaid Services (CMS; formerly known as the Health Care Financing Administration) to reimburse U.S. hospitals for the care they provide to Medicare patients. The RVUs reflect the effect of diagnosis on the relative quantity of resources used in a hospital stay. In addition to DRG weights, the cost function included length of stay, demographic and other clinical information. The method we employed makes the following assumptions: (1) that the non-VA relative value units, the Medicare DRG weights, reflect the relative costs of VA hospital

---

[1] The federal fiscal year begins on October 1 and ends on September 30 of the following year. The convention is to refer to a federal fiscal year (FY) by the year it ends, thus FY98 represents the period October 1, 1997 to September 30, 1998.

stays, and (2) that all stays with the same characteristics have the same cost.

- To find the cost of long-term care, we employed relative value weights known as resource utilization groups (RUG). Therefore, costs of long-term care are adjusted for case-mix as measured by the RUG score. Veterans with higher RUG scores are considered to have higher costs (FY98-00 only).

NEW

- In FY01 & FY02, the cost of long-term care is a per diem rate. In FY01, VA switched from RUG II to the RUG III/MDS dataset. These new RUG scores are not yet available.

1.1 Departures from past years' methods

As the average cost method evolves, improvements are made. Below is a brief summary of the changes that were adopted with the FY98, FY99 and FY00 datasets.

1.1.1 Acute medical-surgical short stay hospitalizations

Prior to FY98, average costs were based on a cost function that estimated an individual's cost as a function of the deviation from the medical center average (Barnett et al., 2000). One problem with this approach is that there was often not enough variation to precisely estimate individual-level costs. Therefore, a new approach was to use a cost function based on Medicare data. We made the cost function's form highly flexible to account for variations in severity and length of stay.

1.1.2 Categories of inpatient care

In prior years, average daily rates were calculated for 10 types of inpatient care (Barnett et al., 2000). Starting in FY98, we added an eleventh category for psychosocial residential rehabilitation programs (PRRTP). PRRTP care was designed to be a less intensive and a less costly type of substance abuse and psychiatric care. A separate time series analysis confirms that PRRTP care is related to decreases in the cost of substance abuse and psychiatric care for those medical centers that have it (Wagner & Chen, 2000).

1.1.3 Nursing home care

In prior years, an average daily rate was calculated for nursing home stays. This method, which reflects case-mix adjustment, was used for FY98-FY00. Using the Resource Utilization Groups (RUGs) that are collected biannually on nursing home patients, we imputed the daily cost per RUG unit. To obtain the patient's cost per stay, we multiplied each patient's rug score by the per rug cost times the length of stay. This is discussed in greater detail in Chapter 6.

In FY01, the cost of long-term care is a per diem rate. In FY01, VA switched from RUG II to the RUG III/MDS dataset. These new RUG scores are not yet available.

1.2 Organization of paper

The average cost estimates represent a merger of centralized VA cost and utilization databases and relative value units obtained from non-VA databases. This paper begins with a description of the VA Cost Distribution Report (CDR), our source of VA cost information. Section 3 covers the utilization data. Chapter 4 provides an overview of our method of merging the CDR with the VA utilization files.

Section 5 describes our method of determining the daily cost of non-medical/surgical care: rehabilitation, blind rehabilitation, spinal cord injury, psychiatry, substance abuse, domiciliary, and intermediate medicine. Chapter 6 describes the methods for estimating the cost of nursing home stays. Chapter 7 describes our method of finding the cost of acute medical-surgical hospital stays. Chapter 8 is a users guide that provides some information on using the data at Austin.

Chapter 2. Cost Distribution Report

The Cost Distribution Report (CDR), also called report RCS 10-0141, is routinely prepared by all VA medical centers. The CDR represents an estimate of the costs expended by each VA patient care department.

VA expenditures are recorded in its general ledger, the Financial Management System (FMS). The FMS system tracks expenditures by cost center, a budget entity that corresponds to a VA service. Examples of VA cost-centers are Medical Service, Nursing Service, and Plant Operations. Cost centers do not correspond to a specific patient care department.

The CDR is created by distributing costs reported in the FMS cost centers to the "cost distribution accounts" (CDA) of the CDR. The CDAs include patient care departments, such as Medical Intensive Care, or Ambulatory Care, Medicine. CDAs also include indirect cost departments.

The distribution of costs is based on estimates prepared by the service chiefs in each medical center. Each service chief estimates the amount of time staff spent on different activities. The cost of staff time, as reported in FMS, is then assigned to each CDA. At the end of each fiscal year, a cumulative CDR is prepared, and it is reconciled to the costs reported in FMS.

Table 2.1 lists the inpatient cost distribution accounts in the CDR. There are additional cost accounts, such as ambulatory care, cost of contract providers, home care programs, and benefits, which are not included in the table.

Table 2.1 also shows the correspondence between direct and indirect costs in the CDR. The middle column lists the direct cost CDAs. These represent costs directly attributed to patient CDAs, such as the cost of physician services, nursing staff, laboratory services, supplies, etc. The right column provides the indirect CDAs. The CDR does not distribute these indirect costs to each department; instead, they are only distributed to a group of departments. While there are more than 40 direct cost accounts, there are just 7 corresponding indirect cost accounts.

Each of these indirect CDA accounts includes as many as eleven different types of indirect costs, each of which is distinguished by numbers to the right of the decimal place. The types of indirect costs include education (.11, .12, .13, .14), research (.21 and .22), administrative support (.30), building management (.40), engineering (.50), equipment depreciation (.70), building depreciation (.80). Thus the indirect cost account "medical research support" for medical bedsection is designated as 1100.21, and includes the costs of medical research associated with the eleven CDAs numbered between 1100 and 1118. We used the CDR detail file as our source of data, as it includes indirect cost CDAs for equipment and building depreciation that are not included in the CDR jurisdictional file.

2.1 Distribution of indirect costs

Our average cost estimate required information about the cost of each category of inpatient care, including its share of indirect costs. The CDR distributes indirect costs only to groups of patient care departments, but we needed to distribute them to each CDA. We assigned indirect costs to each CDA in proportion to its share of the total direct costs of its group of CDAs. For example, the indirect cost of the inpatient mental health sections was distributed to the component departments of psychiatry, substance abuse, and PTSD according to each CDA's share of their total direct cost. At a facility where the psychiatry CDA had 55% of the direct

Table 2.1: Cost Distribution Accounts (CDAs) in the
Cost Distribution Report Inpatient Services (as of FY00)

<table>
<tr><th></th><th colspan="2">Cost distribution account</th></tr>
<tr><td>Department</td><td>Direct cost</td><td>Indirect cost</td></tr>
<tr><td>General medicine</td><td>1110</td><td rowspan="11">1100</td></tr>
<tr><td>Neurology</td><td>1111</td></tr>
<tr><td>Rehabilitation</td><td>1113</td></tr>
<tr><td>Epilepsy center</td><td>1114</td></tr>
<tr><td>Blind rehab</td><td>1115</td></tr>
<tr><td>Spinal cord injury</td><td>1116</td></tr>
<tr><td>Medical ICU</td><td>1117</td></tr>
<tr><td>Inpatient dialysis</td><td>1118</td></tr>
<tr><td>Inpatient aids</td><td>1119</td></tr>
<tr><td>Gem unit - med beds</td><td>1120</td></tr>
<tr><td>Primary care – med</td><td>1130</td></tr>
<tr><td>Surgical ward cost</td><td>1210</td><td rowspan="5">1200</td></tr>
<tr><td>Surgical ICU</td><td>1211</td></tr>
<tr><td>Operating room suite</td><td>1212</td></tr>
<tr><td>Open heart surgery</td><td>1213</td></tr>
<tr><td>Primary care – surg</td><td>1230</td></tr>
<tr><td>Psychiatric wd cost</td><td>1310</td><td rowspan="10">1300</td></tr>
<tr><td>Gen intermediate psych</td><td>1311</td></tr>
<tr><td>S/A intermediate care</td><td>1312</td></tr>
<tr><td>S/A treat program– hi</td><td>1313</td></tr>
<tr><td>Spec inpat PTSD unit</td><td>1314</td></tr>
<tr><td>Eval/brief trmt PTSD</td><td>1315</td></tr>
<tr><td>Star I, II & III</td><td>1316</td></tr>
<tr><td>S/a star I, II & III</td><td>1317</td></tr>
<tr><td>Gem unit - psych bed</td><td>1320</td></tr>
<tr><td>Primary care – psych</td><td>1330</td></tr>
<tr><td>VA nursing home care</td><td>1410</td><td rowspan="2">1400</td></tr>
<tr><td>Gem unit – NH beds</td><td>1420</td></tr>
<tr><td>Domiciliary bed sect</td><td>1510</td><td rowspan="4">1500</td></tr>
<tr><td>Dom substance abuse</td><td>1511</td></tr>
<tr><td>PTSD resid rehab dom</td><td>1512</td></tr>
<tr><td>Homeless domiciliary</td><td>1513</td></tr>
<tr><td>Gem unit – dom beds</td><td>1520</td><td></td></tr>
<tr><td>Intermediate care</td><td>1610</td><td rowspan="2">1600</td></tr>
<tr><td>Gem unit - int beds</td><td>1620</td></tr>
<tr><td>PRRTP</td><td>1711</td><td rowspan="6">1700</td></tr>
<tr><td>PRRP</td><td>1712</td></tr>
<tr><td>SARRTP</td><td>1713</td></tr>
<tr><td>HCMI cwt/tr</td><td>1714</td></tr>
<tr><td>SA cwt/tr</td><td>1715</td></tr>
<tr><td>General cwt/tr</td><td>1717</td></tr>
</table>

cost in the group of inpatient mental health CDAs, we assigned 55% of the indirect cost to psychiatry.

We considered using quantity of utilization as the basis for allocating indirect costs. This would have required us to assume that indirect costs are incurred in proportion to the quantity of service provided, such as the number of inpatient days. We decided that this assumption was problematic, as services are very heterogeneous. For example, since some stays have much greater direct costs, it is not reasonable to assume that they use the same indirect cost. We are unaware of any available data to distribute VA indirect costs on another basis, e.g., to distribute facility maintenance costs based on square footage of space.

2.2 CDR units and unit costs

We did not use the units of service or the unit costs reported in the CDR because of our concerns in the accuracy of these data. We have found that utilization is sometimes excluded. This occurs when a cost distribution account has no cost; any utilization in the corresponding bedsection or clinic stop is not included in the CDR. In addition, costs are sometimes excluded from the calculation of unit costs. This occurs when the CDR reports costs but has no matching utilization, since unit costs would otherwise be a "divide by zero" error, the computer program that creates the CDR calculates the unit costs for that department to be zero; in this way, the cost is effectively dropped from consideration. Rather than use these units or unit costs, we used the VA Patient Treatment Files as our source of utilization data in order to find the per unit cost of services.

Chapter summary

- The CDR represents an estimate of the costs expended by each VA patient care department.
- The CDR is created by distributing costs reported in the FMS cost centers to the "cost distribution accounts" (CDA) of the CDR.
- The CDAs include patient care departments, such as Medical Intensive Care, or Ambulatory Care, Medicine.
- CDAs also include indirect cost departments. The CDR does not distribute these indirect costs to each department; however, they are only distributed to a group of departments.
- We assigned indirect costs to each CDA in proportion to its share of the total direct costs of its group of CDAs.
- We did not use the units of service or the unit costs reported in the CDR because of our lack of confidence in the accuracy of these data. Instead, we used the VA Patient Treatment File as our source of utilization data in order to find the per unit cost of services.

Chapter 3. VA Inpatient Databases

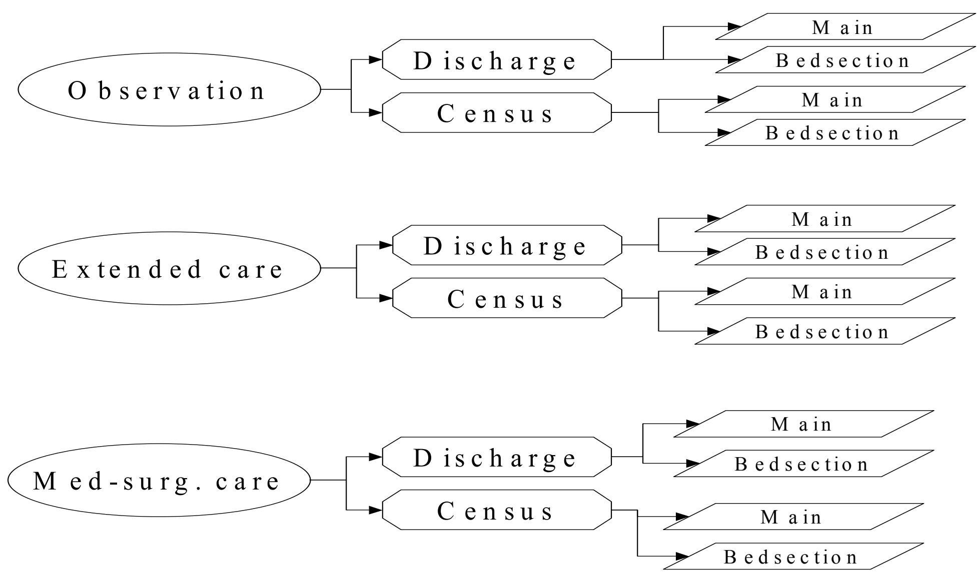
The VA maintains a database of hospital stays called the Patient Treatment File (PTF). Although this database contains neither cost nor charge data, it includes data such as patient demographics, length of stay, and the Diagnosis Related Group (DRG) for the hospitalization.

3.1 VA Utilization Datasets

The PTF records information on hospital stays in different files. It is important to understand how this information is organized because VA defines a hospital stay somewhat differently than non-VA hospitals.

There are three file types of files: observation, extended care and other care. The observation, extended care and other care have a main and a bedsection, and for each of these there is a discharge and census file. As shown in Figure 3.1, there are 12 files.

Figure 3.1: VA inpatient data files



3.1.1 PTF med-surg main discharge file (PTF Main)

This file reports all hospital stays that ended in a particular year. This file contains one record for each hospital stay. The main file does not use a definition of a hospital stay that is strictly comparable to non-VA hospitals. In the non-VA sector, an acute medical-surgical hospitalization followed by a long-term care stay would be recorded as two different stays. In the PTF main file, however, this is often recorded as a single stay.

In some cases, the PTF main is analogous to the non-VA sector. For example, an acute medical-surgical hospital stay that began in the Intensive Care Unit and ended in a medicine ward would be regarded as a single stay in the non-VA sector. This would be recorded as a single record in the PTF main file.

We wanted to apply relative value units from acute medical-surgical stays in non-VA hospitals to estimate the cost of acute medical-surgical VA hospital stays. This required us to develop a definition of what is an acute medical-surgical hospital stay. We used information from both the main and bed-section files to define an acute medical-surgical inpatient stay; see Chapter 7 for a description of our methods for finding the cost of acute medical-surgical hospital stays.

3.1.2 PTF acute care bedsection discharge file (PTF bedsection)

The PTF Bedsection file, is similar to the PTF Main, except that it has multiple records per stay. The PTF Bedsection file divides hospital stays into sequential segments, with one record for each portion of the stay spent in a different bedsection. A bedsection is a hospital ward such as medicine, intensive care, rehabilitation, or long-term care. The bedsection view provides information on the number of days the patient spent in each bedsection.

The PTF Bedsection file does not contain the same data elements as the PTF discharge main file. It is necessary to use both files to obtain all of the hospital discharge information that is required. For example, gender, age and number of diagnoses are in the PTF main discharge file but not in the bedsection discharge file.

There are other slight, but important distinctions between the PTF Main and Bedsection files. As mentioned above, the Main file does not use a definition of a hospital stay that is strictly comparable to the non-VA sector. Both acute medical-surgical stays and non-medical/surgical stays are all aggregated in the PTF Main, while the non-VA sector would typically have an acute medical-surgical stay record and a non-medical/surgical stay record. The PTF Bedsection file, on the other hand, separates stays into each bedsection stay. Hence a stay with an acute medical-surgical bedsection component and a non-medical/surgical component would have two records, which is analogous to the non-VA sector. However, the PTF bedsection file also separates transfers between acute medical-surgical bedsection or between non-medical/surgical bedsections. Such transfers result in more than one record in the PTF Bedsection file. In the non-VA sector, transfers between acute medical-surgical wards would be considered as part of one stay as long as the patient was not transferred to a non-medical/surgical ward during the stay.[2]

3.1.3 PTF Acute Census files

The PTF main and bedsection files include information on all stays that ended during a given fiscal year, regardless of when they began. As is common with discharge files, they do not report on people occupying a bed at the end of the reporting period. To fill this gap, the PTF Census Files includes information on patients who are in a VA hospital at the end of the fiscal year. Note that Census files are given the name of that fiscal year in which they ended. For

---

[2] The bedsection is the "treating specialty" assigned to the physician who is responsible for the patient's care. It roughly corresponds to the location where care is delivered. We used this variable from the PTF, called BEDSECN, to characterize inpatient care. PTF includes another variable, PLBED, to denote the location where care was provided. We did not use this variable to characterize the location of care because many records have missing values for PLBED, whereas all records have a value for BEDSECN.

example, Census FY98 was completed in September 30 1998.[3]

3.1.4 PTF Extended care files

As shown in Figure 3.1, the inpatient utilization files at Austin are divided into three components pertaining to acute inpatient care, extended care, other observation stays. Most stays that start in a nursing home file are included in the extended care file, regardless of the bedsection in which the patients ends up. On the other hand, stays that do not start in the nursing home are usually listed in the non-extended care files, even if the patient was transferred to a long-term care unit.

Since stays may be made up of both acute medical-surgical and long-term care, both of these files contain information on stays that involve acute medical-surgical and long-term care bed-sections. The assignment of stays to one set of files or the other did not affect our treatment of data. We merely used all data from both sets of files for our calculations.

3.1.5 Observation Bed files

The Observation Bed file was first created in 1998 and has been used with increasing frequency in each year since then. If a stay includes an observation bedsection, then the observation portion of the stay is separated from the rest of the stay and included in this file. Most observation bed stays were one day stays, with the patient being discharged from the hospital. However, in some cases there are observation stays that preceded an acute medical-surgical hospital stay. In a few rare instances, people were discharged from an acute medical-surgical hospital stay to the observation bed. In the latter two examples, the portion of the stay that corresponds to the observation bed is kept in the observation bed file.

Observation bedsections were created at the same time as the VA was implementing managerial performance incentives to reduce the number of inpatient days per 1000 treated veterans. Observation data are not included in this performance measure.

Because observation bed stays are so heterogenous, they present some difficulty in determining their cost. For FY98-FY00, we decided that all observation stays would be given the daily cost of the marginal cost of a day. To calculate the marginal cost of day, we used a statistical model with Medicare data (see Chapter 7).

---

[3] The one cautionary note with the Census file is that not all bedsections are coded on September 30 or October 1. Some stays cross the fiscal year are logged in on October 2 and 3. On rare occasions, the stay may be logged in as late as two weeks after. To get an accurate estimate, rather than rely on Census counts for October 1, we recommend that people use the bedsection inday and outday variables to identify whether the person was in a bed at the end of the fiscal year.

Chapter summary

- The Patient Treatment File (PTF) records information on hospital stays are in two different datasets (PTF main and PTF bedsection).

- The PTF main file reports all hospital stays that ended in a particular year.

- The PTF utilization files (Main and Bedsection) are divided into three components pertaining to acute care, extended care, observation stays. Acute care, extended care and observation stays each have a discharge and a census file.

- The bedsection file divides hospital stays into sequential segments, with one record for each portion of the stay spent in a different bedsection. A bedsection is a hospital ward such as medicine, intensive care, rehabilitation, or long-term care.

- The PTF Main and Bedsection are discharge files and they do not report on people currently occupying a bed at the end of the fiscal year. To fill this gap, the PTF Census Files includes information on patients who are in a VA hospital at the end of the fiscal year.

Chapter 4. Merger of cost and utilization databases

This section describes how we merged the CDR with VA utilization databases. The VA database of hospital stays is called the Patient Treatment File (PTF). This paper does not cover outpatient data.

We excluded the cost of facilities that do not provide patient care. In addition, we accounted for mergers between VA medical centers. Over time, facilities have consolidated, but these consolidations were not necessarily implemented at the same time in the cost and utilization databases. We recoded data to keep a common definition of facility in the databases. Since patient care departments are sometimes defined differently in the cost data than in the utilization data, we aggregated departments to find a common denominator.

4.1 Excluded facilities

We excluded the 16 facilities that report costs in the CDR, but do not report utilization in either the PTF or the OPC. These include records for VA Headquarters (station 101), information services centers, and other VA support facilities. A list of these facilities, and their three digit facility number, is provided in Table 4.1. Nine of these facilities do not appear in the official listing of VA facilities.[4]

Table 4.1: Excluded Facilities

Facility Number	Facility Name
101	VHA Headquarters
200	Austin Automation Center
722	Albuquerque, NM Outpatient Center
741	Denver CHAMPVA
721, 724, 742, 760, 761, 762, 763, 764, 765	
792	Prosthetics Center
794	Somerville
797	Hines (CIO)

We felt that central administration may involve activities that are more characteristic of a health care payer, rather than a health care provider. For this reason, we decided not to count these facility's costs as overhead costs that should be distributed to patient care departments.

4.2 Facility mergers

VA has been consolidating facilities. When one facility merges with another, they both take on a single identification number (see Table 4.2). This change is sometimes implemented at different times in the different data systems. We wished to maintain the distinction between facilities as long as it was possible. We also wished to work with observations that consisted of

[4] Consolidated Address and Territorial Bulletin 1-L, U.S. Department of Veterans Affairs, Washington, DC 20420, August 31, 1999

facility level data for an entire fiscal year. We consolidated all data into the new facility number in the first fiscal year that the CDR or the utilization databases no longer maintained a distinction between the facilities.

Table 4.2: Facility Consolidations in 1997-2000

	Old STA3N	New STA3N
1997		
VA Chicago Health Care System (Westside/Lakeside)	535	537
VA Central Alabama Health Care System (Montgomery/Tuskegee)	680	619
VA North Texas Health Care System (Dallas/Bonham)	522	549
Southern California System of Clinics (Sepulveda/Los Angeles/Santa Barbara)	665,752	665
Hudson Valley VA Health Care System (Montrose/Castle Point)	533	620
VA Central Iowa Health Care System (Des Moines/Knoxville)	592	555
VA Greater Nebraska Health Care System (Lincoln/Grand Island)	574	597
1998		
VA Eastern Kansas Health Care System (Topeka/Leavenworth)	686	677
VA Montana Health Care System (Fort Harrison/Miles City)	617	436
North Florida/South Georgia Veterans Health Care System (Gainesville/Lake City)	594	573
VA Greater Los Angeles Health Care System (West Los Angeles/Southern California System of Clinics)	752	691
1999		
North Florida/South Georgia HCS	594	573
Greater Los Angeles HCS	665	691
Boston VAMC	525	523
2000		
NY Harbor Health Care System (Brooklyn Poly Pl)	527	630
Upstate NY Health Care System (Canandaigua)	532	528
Upstate NY Health Care System (Shreveport)	670	528
VA Mid Tennessee Health Care System (Murfreesboro)	622	626
Upstate NY Health Care System (Albany)	500	528
VA Nebraska Western (Iowa City)	584	636
2001		
Columbia MO	543	589
Topeka KS	677	589
Marion IL	609	657
Popular Blue MO	647	657

4.3 Definition of patient care unit

Patient care units are defined differently in the CDR and the utilization databases. In the CDR, care is characterized by the Cost Distribution Account. The Cost Distribution Report Handbook maps the correspondence between Cost Distribution Accounts and the utilization databases. It does not include the Cost Distribution Accounts and utilization codes created since 1996, so the handbook is now out of date.

The Patient Treatment File (PTF) characterizes inpatient care by bedsection, which refers to the ward where the patient received care, such as medical intensive care unit, or nursing home unit. Each inpatient Cost Distribution Account in the CDR reports the costs of operating a group of several different bedsections. To learn about the correspondence between new bedsection codes and new cost distribution accounts, we examined the variable BEDCDR in the PTF bedsection file. This variable has the value of the CDA that corresponds to the bedsection. Only one CDA is assigned to each bedsection. As a result, the exact correspondence between BEDCDR and BEDSECN (the variable for bedsection) in the PTF represents a statement of the CDA associated with each bedsection.

Our review of CDR data suggests that some medical centers do not consistently use the definitions given in the CDR handbook and in these supplemental sources. The cost of providing care in a particular bedsection is not always assigned to the corresponding CDA specified in the CDR handbook. Some facilities have utilization in bedsections without assigning any costs to the corresponding CDA. In other cases, costs are assigned to a CDA, but no utilization appears in the corresponding bedsections.

One cause of this problem is the addition of new CDAs to the CDR and new bedsections to the PTF. Facilities may implement new utilization codes and CDAs at different times.

We dealt with these potential problems by defining aggregate “patient care categories.” These categories represent our best judgment about what constitutes the smallest common denominator between cost and utilization. A patient care category represents a group of related cost distribution accounts, and their associated utilization.

We defined patient care categories based on earlier work (Barnett et al., 2000). We aggregated CDAs into eleven categories, and ascertained that for almost every medical center, if the category had costs, it also had utilization, and if it had utilization, it also had costs. We also examined the mean cost of care, examining outliers that suggested mismatch of cost and utilization data.

For some categories of care at some medical centers, there were mismatches between cost and utilization data. Most mismatches were handled by assigning the costs and utilization to a similar department, creating a higher level of data aggregation. For more details on the reconciliation, see 4.5.

4.4 Merger of cost and inpatient utilization data

The CDR reports on expenditures in a federal fiscal year, which runs from October 1 until September 30. As mentioned above, we wanted to identify the amount of care provided during the fiscal year. Since hospital stays may span fiscal years, we developed a method to divide hospital utilization between fiscal years.

The denominator for the cost data was the fiscal year, whereas the denominator for the utilization data was discharges. These denominators are not equivalent. We could have ignored this difference. This would have been equivalent to assuming that bed occupancy was constant over the year. However, this assumption would be wrong because we know that there is a trend to shorten length of stay and to reduce hospitalization. And we did not want to assume that the same number of patients is in the hospital at the start and at the end of the fiscal year.

A better way to adjust for the difference in denominators was to use information from the Census files. With the Census files we adjusted the discharge file so that it more closely approximated utilization in the fiscal year.

For the utilization data, we included days spent during the current fiscal year by all patients. For those discharged during the fiscal year, their data came from the PTF, limiting the days to those in the fiscal year. For those patients not discharged by the end of the fiscal year, we obtained these days of stay from the PTF census files. This calculation included "leave" days, that is, days that a patient was absent from a hospital, though not yet discharged.[5] The PTF records leave days, but it does not indicate when they occurred. We assumed that leave days are uniformly distributed throughout the stay.

The finest level of detail for the cost data is at department level; patient-level cost data do not exist. To merge the cost and utilization data, we identified 11 categories of inpatient care (see Table 4.3).

Before FY98, we used 10 categories to characterize all inpatient care. Starting with FY98, we created an 11th category for Psychosocial Residential Rehabilitation Programs (PRRTP). PRRTP programs are less intensive inpatient programs for psychiatry and substance abuse. A separate time series analysis confirms that medical centers that adopted PRRTP care had an associated decrease in the cost of substance abuse and psychiatric care (Wagner & Chen, 2000).

Table 4.3: Categories of Inpatient Care **Table 4.3: Categories of Inpatient Care**

Category of Care	CDR acct	BEDSECN
Inpatient Medicine	1100*,1110, 1114,1117, 1118,1119, 1120,1130	1-12,14-17,18[#], 19,24[#], 31,34,35,75,83
Inpatient Rehabilitation	1100*,1113	20, 41[#]
Inpatient Blind Rehabilitation	1100*,1115	21, 36[#]
Inpatient Spinal Cord	1100*,1116	22, 23[#]
Inpatient Surgery	1200-1213, 1230	50-63, 65[#]
Inpatient Psychiatry	1310,1311,1314,1315,1316,1317,1320, 1330, 1711[^],1712[^],1714[^],1717[^]	25[^],26[^],28[^],33,70,71,76,77,79, 89,91,92,93,94[#]
Inpatient Substance Abuse	1312,1313, 1713,[^] 1715[^]	27[^],29[^],72-74,84,90
Inpatient Intermediate	1600,16101620	32,40
Inpatient Domiciliary	1500,1510,1511,1512,1520, 1513^	37^,85-88
Inpatient Long Term	1400,14101420	80,81
PRRTP[^]	1711, 1712, 1713, 1714, 1715	25,26,27,28,29, 38[#], 39[#]

"*" indicates an overhead account whose costs were allocated proportionally across more than one care category.
[^] These CDR accounts and bedsections were assigned to psychiatry and substance abuse at medical centers that did not have an official PRRTP program. In FY00 PRRTP programs existed at: 501, 504, 463, 637, 515, 516, 518, 523, 528, 541, 549, 554, 561, 568, 573, 590, 459, 586, 589, 555, 595, 598, 546, 620, 622, 556, 631, 632, 635, 640, 645, 653, 658, 662, 663, 666, 656, 676, 678, 687, 689
^ New for FY00
[#] New for FY01

[5] Leave days are also called Absent Bed Occupant Days and are given the variable name LVB in the PTF

4.5 Data reconciliation

After using the 11 inpatient categories to merge the cost and utilization data for each medical center, we performed a reconciliation. This was necessary because the VA does not routinely reconcile these two databases. The most obvious discrepancies are when a category has costs but no utilization. The opposite can also be true– utilization exists without costs. In reality, the occurrence of these discrepancies is quite rare. When they occurred we merged the substance abuse costs and utilization with the psychiatry costs and utilization. Appendix A describes all the reconciliations that were done for FY98-FY01.

4.6 Daily rate

After reconciling the 11 inpatient categories, there was a direct correspondence between costs and utilization. We divided total costs by total utilization to find the average cost for each category of care at each medical center. We compared rates across medical centers, and we found the average rate for each of the categories. Table 4.3 lists the average rates for inpatient care in FY98 -FY01.

It is important to note that this daily rate does not account for case mix, clinical information or demographic characteristics. It is just an average daily rate. To use these rates, one would have to assume that costs are only a function of length of stay. This is not an extremely appealing assumption. Unfortunately, for most of the categories we have little additional information that can be used to make more accurate cost estimates. For acute medicine and surgery, we have a better method for estimating costs, which is covered in Chapter 7. For nursing home care, we have developed a new method that accounts for case-mix; this is presented in Chapter 6.

Table 4.4: Median facility cost per day of stay for inpatient care, FY98-01

	Category of Care	FY98	FY99	FY00	FY01	FY02
1	Inpatient Medicine*	$1,195	$1,304	$1,319	$1,381	$1,465
2	Inpatient Rehabilitation	$890	$1,029	$1,012	$1,102	$1,377
3	Inpatient Blind Rehabilitation	$728	$762	$815	$834	$861
4	Inpatient Spinal Cord	$764	$838	$791	$843	$971
5	Inpatient Surgery*	$2,625	$2,797	$2,455	$2,700	$2,882
6	Inpatient Psychiatry	$680	$745	$744	$769	$864
7	Inpatient Substance Abuse	$821	$576	$418	$595	$666
8	Inpatient Intermediate	$625	$548	$525	$599	$794
9	Inpatient Domiciliary	$126	$238	$126	$162	$173
10	Inpatient Long Term	$275	$303	$305	$339	$358
11	PRRTP	$161	$179	$179	$213	$220

Includes overhead costs

* We do not recommend using this daily rate as we have provided more accurate estimates (see Chapter 7).

Chapter summary

- We excluded the cost of 16 facilities that do not provide patient care. We felt that central administration may involve activities that are more characteristic of a health care payer, rather than a health care provider.

- We also accounted for mergers between medical centers. If medical centers merged during a fiscal year, we merged their utilization and cost data for the entire fiscal year. It was not possible to separate costs and utilization before and after the merger.

- Patient care units are defined differently in the CDR and the utilization databases. In the CDR, care is characterized by the cost distribution account. The Patient Treatment File (PTF) characterizes inpatient care by the "bedsection."

- Our review of CDR data suggested that many medical centers do not consistently use the definitions given in the CDR handbook. We dealt with this by defining aggregate 11 "patient care categories."

- In merging the PTF data to the CDR data, one must remember that the PTF has a discharge view while the CDR takes a fiscal year view. These are not synonymous views and an adjustment is needed to make these equivalent.

- Even for these patient care categories there was not always a one to one correspondence between the CDR and the PTF. We did our own reconciliation to solve this problem. The exact reconciliations are provided in an Appendix A.

- After reconciling the 11 inpatient care categories, we generated an average cost per day in each category.

Chapter 5. The cost of non-medical/surgical inpatient care

5.1 What is non-medical/surgical inpatient care?

Most hospitals in the US differentiate between short-stay acute medical-surgical and non-medical/surgical hospitalizations. Short-stay acute medical-surgical hospitalizations are generally for acute medicine and surgical treatment. While over 90% of short stay hospitalizations are less than 60 days long, there are rare cases that involve a length of stay up to and over a year. In the VA, 88% of inpatient stays can be categorized as acute medical-surgical; in other words, we can show that 88% of the stays fit in acute bedsections (defined as medicine and surgical, see Table 3.1). The remaining 12% of stays are non-medical/surgical; this includes to rehabilitation, blind rehabilitation, spinal cord injury, psychiatry, substance abuse, intermediate care, domiciliary, and nursing home. This chapter describes how we estimated the cost for non-medical/surgical care for FY98-FY01. The one difference between FY98 and prior years is the use of case-mix adjustment for nursing home care. The cost of nursing home care is covered in Chapter 6.

5.2 Cost methodology for non-medical/surgical care

Determining costs for non-medical/surgical care is the most straightforward of the cost determination methods. The premise is to merge the CDR and PTF bedsection databases for each of the 11 care categories. The 11 care categories are defined by bedsection and cost distribution accounts (see Table 3.1). Two values are needed to calculate a daily cost for each of the care categories: total costs and total number of days. With this information, a daily rate can easily be calculated by dividing total costs by total days. This can be done either at the medical center level or for the entire nation. When this is done at the level of the medical center, the result is an average daily rate for that medical center. We refer to this rate as the local daily cost estimate.

5.2.1 Leave and pass days

For stays that began before the beginning of the fiscal year, we found the length of stay during the current fiscal year by finding the number of days between the discharge date and the beginning of the fiscal year. This calculation considered "leave" days, that is, days that the patient was absent from the hospital, though not yet discharged. Leave days are also called Absent Bed Occupant Days and are given the variable name LVB in the PTF. The PTF records leave days in a variable named LVB, but it does not record when they occurred. We assumed that leave days are uniformly distributed throughout the stay.

5.2.2 Local outlier costs

As one might expect, there is more variation in the local daily rates than the national daily rates. This raises the question about the accuracy of the local rate. To help identify inaccurate local costs, we generated a flag if a medical center had a daily rate that ± 2 standard deviations from the average of all VA medical centers (for that particular care category). Part of this variation could be explained by factors such as wages. However, some of this variation is due to accounting mistakes or inconsistencies. Therefore, the flag variable allows the analyst to check for outliers when using the local cost estimates.

5.2.3 Why local rates at all?

Given that there is more variation in the local rates than the national rates, one may ask why do we calculate local rates at all. The answer is that sometimes the variation in the local rates is important. Wages are one factor that affects costs, as they depend on the labor market in different geographic localities. If a researcher is interested in the effect of an intervention on a local medical center or VISN, then the local rates may be more appropriate because they partly reflect the wage differentials and other local differences.

5.2.4 Adjusting for case-mix

Although DRGs have been created for mental health and rehabilitation stays, the cost of stays assigned to these DRGs is highly variable. Because DRGs do not explain the variation in cost of rehabilitation and mental health stays, facilities that provide this sort of care were exempted from the Prospective Payment System of Medicare. We estimated the cost of this type of care using the average daily cost.

Chapter summary

- 12% of inpatient VA stays are non-medical/surgical. We categorize non-medical/surgical care into nine categories: to rehabilitation, blind rehabilitation, spinal cord injury, psychiatry, substance abuse, intermediate care, domiciliary, and nursing home.
- Except for nursing home care, our cost methodology is to generate an average daily rate for each category.
- The average daily rate was estimated for each medical center, providing a local cost estimate, or at the national level, providing a national cost estimate.
- As one might expect, there is more variation in the local daily rates than the national daily rates.
- We generated a flag if a medical center had a daily rate that was ± 2 standard deviations from the average of all VA medical centers (for that particular care category). Part of this variation could be explained by factors such as wages. However, some of this variation is due to accounting mistakes or inconsistencies. Therefore, one should be informed and check for outliers when using the local cost estimates.

Chapter 6. The cost of nursing home care

VA long-term care patients are evaluated using the Resource Utilization Group (RUG) assessment method. These assessments are performed at admission and twice a year (April and October). In the assessment, a wage-weighted work unit (WWU) is assigned to the patient. The WWU represents an estimate of the relative quantity of resources used to care for long-term care patients (Schneider, Fries, Foley, Desmond, & Gormley, 1988). Starting in FY98, we used the relative values from the RUG assessments to adjust VA long-term care costs for case-mix.

This section describes the methods using numbers from FY98. The methods are the same for FY99 and FY00, although the numbers are different.

NEW In FY01 & FY02, the cost of long-term care is a per diem rate. In FY01, VA switched from RUG II to the RUG III/MDS dataset. These new RUG scores are not yet available.

6.1 Case mix index

In FY98, there were 45,694 nursing home stays in the VA utilization files. To adjust nursing home costs for case mix, we calculated three case-mix indexes:

(1) Patient level case-mix index
(2) Medical center nursing home case-mix index, which is a mean index for all patients at a medical center weighted by the length of stay
(3) National nursing home case-mix index, which is a weighted mean index of VA nursing home patients

To estimate costs that occurred within FY98, we included only the number of days from October 1, 1997 through September 30, 1998. However, to calculate patient case-mix, we included all possible assessments associated with nursing home stays in FY98 from the following six files:

(1) FY972: the admission assessment file in the second half of FY97
(2) FY981 the admission assessment file in the first year of FY98
(3) FY982 the admission assessment file in the second half of FY98
(4) OCT97 the regular assessment file in October 1997
(5) APR98 the regular assessment file in April 1998
(6) OCT98 the regular assessment file in October 1998

The October 1998 assessment was included because it was the best measure of resource use at the end of the fiscal year.

6.2 Patient level case-mix

The Resource Utilization Group (RUG II) instrument contains 17 levels of resource use in six categories (see table 6.1). In each category, a letter indicates different level of resource use. For example, a patient in Rehabilitation B is assigned 1000 for WWU. In this report, the value of WWU is called RUG score. This assessment information is contained in the Patient Assessment File (PAF) at Austin Automation Center.

Depending on the date of admission and the length of stay, the number of assessments that a patient could obtain during a nursing home bed-section stay varies. In general, every nursing home patient is assessed at admission. We calculated an average RUG score weighted by the number of days between assessments as the case-mix index of a nursing home stay. To calculate an average RUG score for resource use, we were concerned that there was no measure of resource use at discharge. Resource use could change substantially at discharge from the last assessment, especially when the patient died at discharge. Therefore, we developed a regression model to estimate a RUG score at discharge.

Table 6.1: RUG II classification and Wage-Weighted Work Units

RUG Category		WWU
Rehabilitation	A	896
	B	1000
Special Care	A	867
	B	976
Clinically Complex	A	484
	B	711
	C	778
	D	929
Behavioral	A	479
	B	640
	C	744
Physical	A	413
	B	546
	C	640
	D	707
	E	820
CHR VENT DEP		1800

6.2.1 RUG score at discharge

To estimate resource use at discharge, we developed two regression models: a one-assessment and a two-assessment model, depending on the number of assessments per patient. People in the one-point model had one previous assessment, whereas patients included in the two-point model had at least two assessments.

For the models, we selected a sample of nursing home discharges that occurred in either October or April; that is the discharges were within 30 days of the last assessment. Selected cases also had at least three assessments during the stay from all assessment files between fiscal year 1994 and 1999.

We used the last assessment as the dependent variable. The explanatory variables included one or two RUG scores from previous assessment(s), discharge status (died in hospital or alive at discharge), and length of stay. The time between two regular assessments is 180 days. Consequently, we gave an estimated RUG discharge score for those nursing home stays in which the last assessment was more than 90 days (half of the length between two regular assessments) before the discharge. If a patient only had one RUG score, we estimated the discharge assessment based on the one-point model. Otherwise, for nursing home stays with more than two RUG scores, we used the two-point model. In a sensitivity analysis (not shown), we also examined models including more than two RUG scores. The coefficients of RUG scores with more than two lag periods were not statistically significant. The two models (one-point and two-point) are specified below.

Two-point model: patient had at least two assessments and discharge was more than 90 before discharge

$WWU_d = b_0 + b_1WWU_1 + b_2WWU_2 + b_3D_1 + b_4D_2 + b_5(WWU_1*D_1) + b_6 (WWU_2*D_1) + b_7Died + b_8LOS240$ ($R^2 = .2940$)

where

$WWU_{d =}$ Estimated RUG score within 30 days of discharge

WWU_1 and WWU_2 = the last two WWU assessment scores (WWU_1 is the most recent assessment)

D_1 = an indicator (D_1 =1 when $WWU_1 - WWU_2 > 0$)

$D_{2 =}$ an indicator (D_2 =1 when $WWU_1 - WWU_2 < 0$)

Died: an indicator (Died =1 when a patient died at discharge)

LOS240: an indicator for length of stay (LOS240 =1 when the length of stay is less than 240 days)

One point model: patient had only one assessment and the assessment was more than 90 days before discharge.

$WWU_d = b_0 + b_1 WWU_1 + b_2$ Died ($R^2 = .2411$)

6.2.2 Average WWU

Based on the admission date and length of stay, patients could obtain different numbers of assessments during a single nursing home stay. Figure 6.1 lists 8 possible combinations of admission and discharge time for a stay within a fiscal year.

An average RUG score (WWU) was calculated based on available assessments using the following formulas for each of the 8 situations in Figure 6.1. It was weighted by the proportion of stay preceded or followed each assessment.

Situation 1

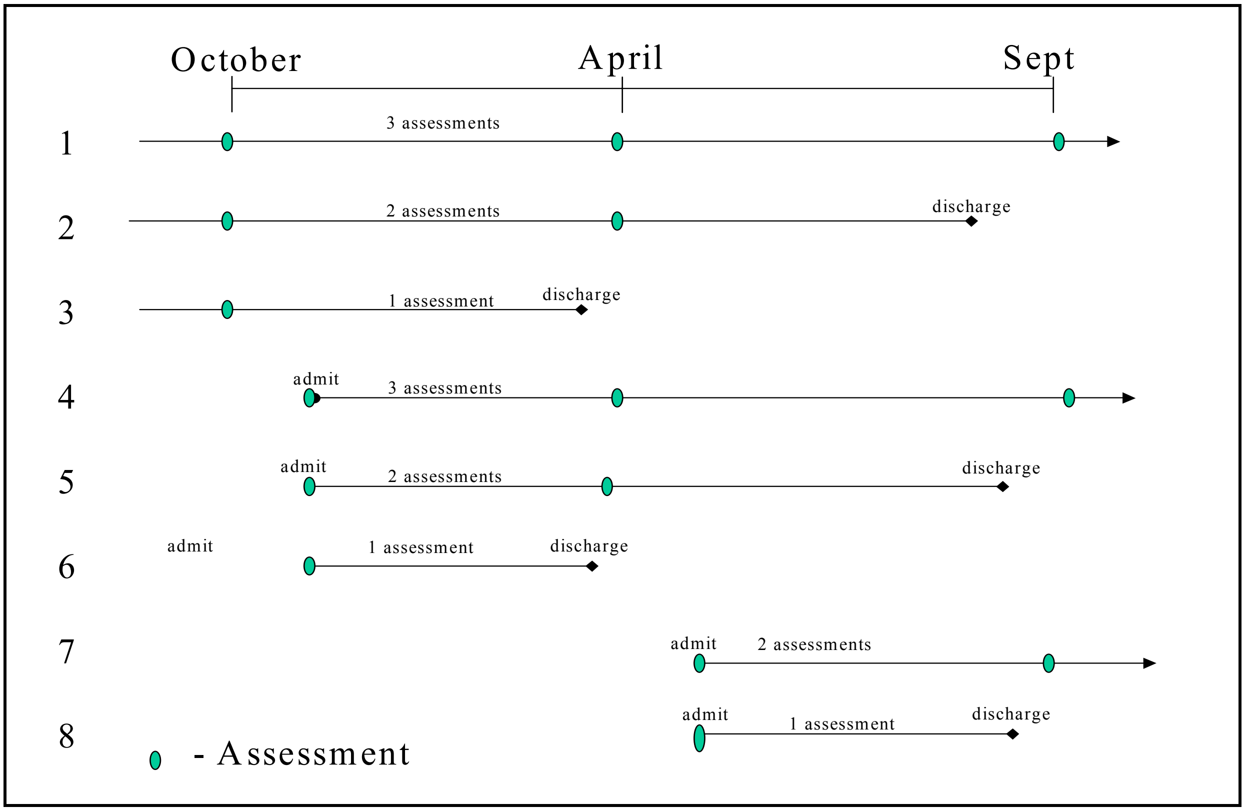
$$\overline{wwu} = \frac{1}{4} wwu_{oct97} + \frac{1}{2} wwu_{apr98} + \frac{1}{4} wwu_{oct98}$$

Situation 2

If the discharge month was between April and June, then

$$\overline{wwu} = \frac{90}{LS} wwu_{Oct97} + \frac{(LS-90)}{LS} wwu_{Apr98}$$

Figure 6.1: Number of possible assessments used to calculate an average WWU in 8 situations



If the discharge month was between July and September, a RUG score (WWU_d) at discharge was estimated using a regression model and the average WWU score was

$$\overline{wwu} = \frac{90}{LS} wwu_{Oct97} + \frac{(90 + \frac{1}{2} LA)}{LS} wwu_{Apr98} + \frac{LA}{2LS} wwu_{Dis2}$$

LS = the number of days from October 1, 1997 through discharge
LA = the number of days from April 1, 1998 through discharge
WWU_{Dis2} = A RUG score estimated by the two-point model.

Situation 3
If the discharge month was between October and December of 1997, the average RUG score (WWU) was the October 1997 assessment.

If the discharge month was between January and March of 1998,

$$\overline{wwu} = \frac{1}{2}(wwu_{Oct97} + wwu_{\text{discharge}})$$

Situation 4

$$\overline{WWU} = \frac{LA}{2LS} WWU_{Adm} + \frac{(90 + \frac{1}{2} LA)}{LS} WWU_{apr98} + \frac{90}{LS} WWU_{Oct98}$$

where
LS = the number of days from admission through September 30, 1998.
LA = the number of days from admission through March 31, 1998.
WWU_{Adm} = RUG score at admission.

Situation 5
If discharge was between April and June of 1998, then

$$\overline{WWU} = \frac{LA_1}{2LS} WWU_{Adm} + \frac{(\frac{1}{2} LA_1 + LA_2)}{LS} WWU_{apr98}$$

where
LS = the number of days from admission through discharge,
LA_1 = the number of days from admission through March 31, 1998, and
LA_2 = the number of days from April 1, 1998 through discharge.

If the discharge month was between July and September of 1998, then

$$\overline{WWU} = \frac{LA_1}{2LS} WWU_{Adm} + \frac{(LA_1 + LA_2)}{2LS} WWU_{apr98} + \frac{LA_2}{2LS} WWU_{Dis2}$$

where
LS = the number of days from admission through discharge,
LA_1 = the number of days from admission through March 31, 1998, and
LA_2 = the number of days from July 1, 1998 through discharge.
WWU_{Dis2} = A RUG score estimated by the two-point model.

Situation 6

$$\overline{WWU} = \frac{1}{2}(WWU_{Adm} + WWU_{Dis1})$$

where WWU_{Dis1} was estimated by the one-point model.

Situation 7

$$\overline{WWW} = \frac{1}{2}(WWU_{Adm} + WWU_{Oct98})$$

Situation 8

Same as the formula used for situation 6.

6.2.3 Exceptions

Among the 45,694 nursing home stays in FY98, 891 did not have any assessments from the 6 assessment files we selected. We assigned the medical center average case-mix as the case-mix index for those nursing home stays.

Among the 44,803 nursing home stays with at least one assessment, 1,432 (3%) did not match with assessments within the expected time windows, which starts from 5 days before and 15 days after the admission date. For these nursing home stays, we calculated an average of up to 3 most recent assessments in FY98 as the case-mix index.

6.3 Case-mix index of a medical center

We calculated a case-mix index for each medical center (LWWU) to measure the average case mix of nursing home patients in the medical center. The LWWU is equal to the sum of case-mix adjusted number of nursing home days divided by the total number of nursing home days in the medical center.

$$LWWU = \frac{\sum (\overline{wwu_i} * LOS_i)}{\sum LOS_i} \qquad (i = 1 \ldots n)$$

where WWU_i is the case-mix index for patient I, LOS_i is the length of stay for patient I, and n is the total number of nursing home admissions in the medical center.

6.4 National case-mix index

We also calculated a national average case-mix index (NWWU) by a similar method:

$$NWWU = \frac{\sum (\overline{wwu_i} * LOS_i)}{\sum LOS_i} \qquad (i = 1 \ldots N)$$

where WWU_i is the case-mix index for patient I, LOS_i is the length of stay for patient I, and N is total number of nursing home admissions in all VA nursing homes during FY98.

6.5 Relative Value Unit (RVU)

Case-mix indexes were normalized at the national as well as at the local (medical center) levels. For each nursing home stay, a national RVU (RVUN) was calculated by dividing the individual case-mix index (WWU) by the national average case-mix index (NWWU) and a local RVU (RVUL) was calculated by dividing the individual case-mix index (WWU) by the local average case-mix index (LWWU). The average case-mix adjusted cost was calculated at two levels: the local (medical center) average cost and the national average cost.

6.5.1 Average case-mix adjusted local cost

The average case-mix adjusted local nursing home cost for patient I at the medical center j was calculated by

$LC_{ji} = DC_j \times LOS_{ji} \times RVUL_{ji}$

where:

LC_{ji} - average case-mix adjusted local nursing home cost for patient I at the medical center j,

DC_j - average non-adjusted average per diem cost of the medical center j,

LOS_{ji} - the length of stay for patient I at the medical center j,

$RVUL_{ji}$ - the local RVU for patient I at the medical center j.

6.5.2 Average case-mix-adjusted national cost

The average case-mix-adjusted national cost was calculated as

$NC_i = DC \times LOS_i \times RVUN_i$

where

NC_i - average case-mix adjusted national nursing home cost for patient I,

DC - average non-adjusted national per diem cost,

LOS_i - the length of stay for patient I,

$RVUN_i$ - the national RVU for patient I.

6.6 Distribution of case-mix

The individual RVUs and the medical center average normalized case-mix indexes for FY98 are listed in table 6.2.

Table 6.2: Distribution of RVUs at Patient and Institutional Levels in FY98

Mean	Std	Min	Max
Individuals			
1.0	0.274	0.59	2.74
Medical Center Means			
1.0	0.094	0.82	1.31

Table 6.2 shows that there is a substantial variation in patient case-mix. The maximum RVU is more than 4 times of the minimum. If nursing home costs were not adjusted for case mix, such large differences in resource use would be missed. Also, the average case-mix for medical centers varies considerably. This could be caused by the differences in patients' health status, institutional characteristics, or the quality of assessment measures. Further investigation is needed to understand these patterns.

Chapter summary

- VA long-term care patients are evaluated using the Resource Utilization Group (RUG-II) assessment method.

- These assessments are performed at admission and twice a year (April and October).

- The assessment assigns Wage-Weighted Work Units (WWU) to the patient. The Wage-Weighted Work Unit represents an estimate of the relative quantity of resources used to care for long-term care patients.

- When a patient has more than one assessment, we calculated a weighted average WWU, with weights reflecting the proportion of the stay that proceeded and /or followed each assessment.

- When the most recent assessment was longer than 90 days from the discharge, we estimated a WWU at discharge using a regression model.

- When a nursing home stay did not have any assessment recorded in the Patient Assessment File, we assigned the institutional average RVU to the stay.

- We used the RUG scores (WWUs) to adjust for resource use. This was done by summing together the number of weighted days for patient stays in a medical center. The total cost from the CDR was then divided by the total number of weighted days, yielding a weighted daily cost. To estimate a person's average cost for a stay, we multiplied the daily cost per weighted day by the weight (RUG score) and the length of stay.

- NEW In FY01 & FY02, the cost of long-term care is a per diem rate. In FY01, VA switched from RUG II to the RUG III/MDS dataset. These new RUG scores are not yet available.

Chapter 7. The cost of acute medical-surgical hospitalizations

The cost of acute medical-surgical hospital care in VA can be more accurately estimated by incorporating diagnostic information from the administrative record (Barnett, 1997). Such methods avoid the assumption that every day of stay is of equal cost. We used an econometric cost function, with parameters estimated from non-VA data, to impute the costs for acute medical-surgical stays in the VA.

This method relies heavily on non-VA relative value weights. These weights, known as DRG weights, are used to pay hospitals for providing care to Medicare patients. Upon discharge, patients are assigned a Diagnosis Related Groups (DRGs) based on their primary diagnosis. This weighting system is used by the Centers for Medicare and Medicaid Services (formerly the Health Care Financing Administration) to determine Medicare payments to hospitals.

This section presents the cost function that we developed with Medicare data. Given the complexities in this chapter, a flow diagram is provided in Appendix B to help readers visualize the process.

7.1 Making an acute medical-surgical inpatient discharge database

The VA keeps track of bedsections (note: bedsection is a VA-specific term that is most analogous to a hospital ward). Because a patient can get transferred among bedsections multiple times within a single acute medical-surgical hospital stay, keeping track of bedsections provides us with a great amount of detail that is necessary for identifying acute medical-surgical stays.

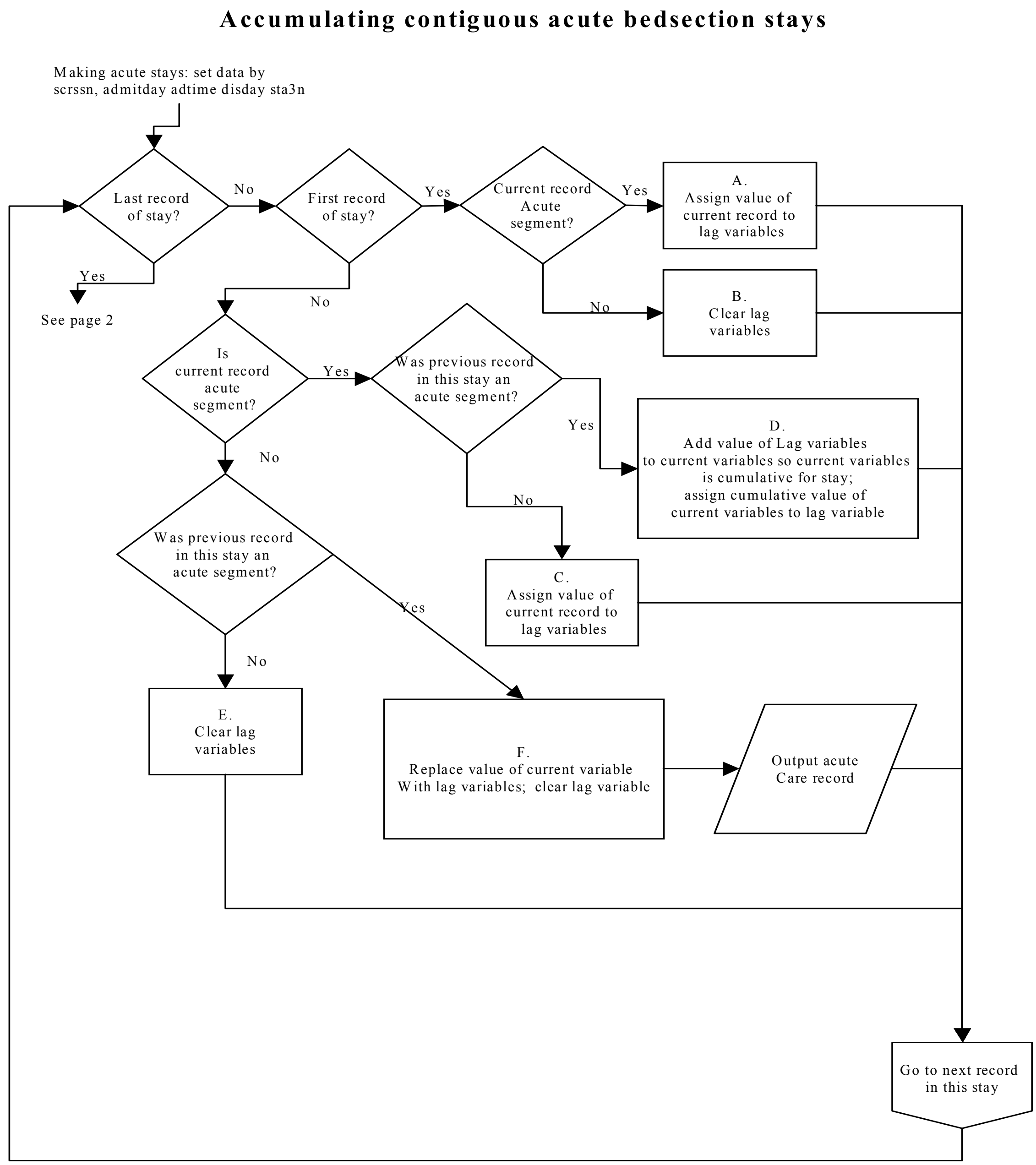
To use non-VA relative value units, we had to restructure the VA data to use the same definition of acute stays as is found outside the VA. Most non-VA databases are organized as discharge databases with each record representing an acute medical-surgical hospital discharge. While the PTF Main is a discharge database, it does not distinguish between acute medical-surgical and non-medical/surgical care. In addition, the PTF Bedsection file is a discharge file but it separates each record into bedsection stays, even if the bedsections are all part of one acute medical-surgical stay. Therefore, we had to make a database of acute medical-surgical discharges using the PTF bedsection file.

We defined an acute medical-surgical stay based on the following bedsections: 01-12, 14-17, 19, 31, 34, 35, 50-63, 75, 83. Of these, the surgical bedsections are 50-63 and the remainder are acute medicine bedsections. These are the bedsections identified by the VA as the source of workload for costs reported in the acute medical and surgical cost distribution accounts.

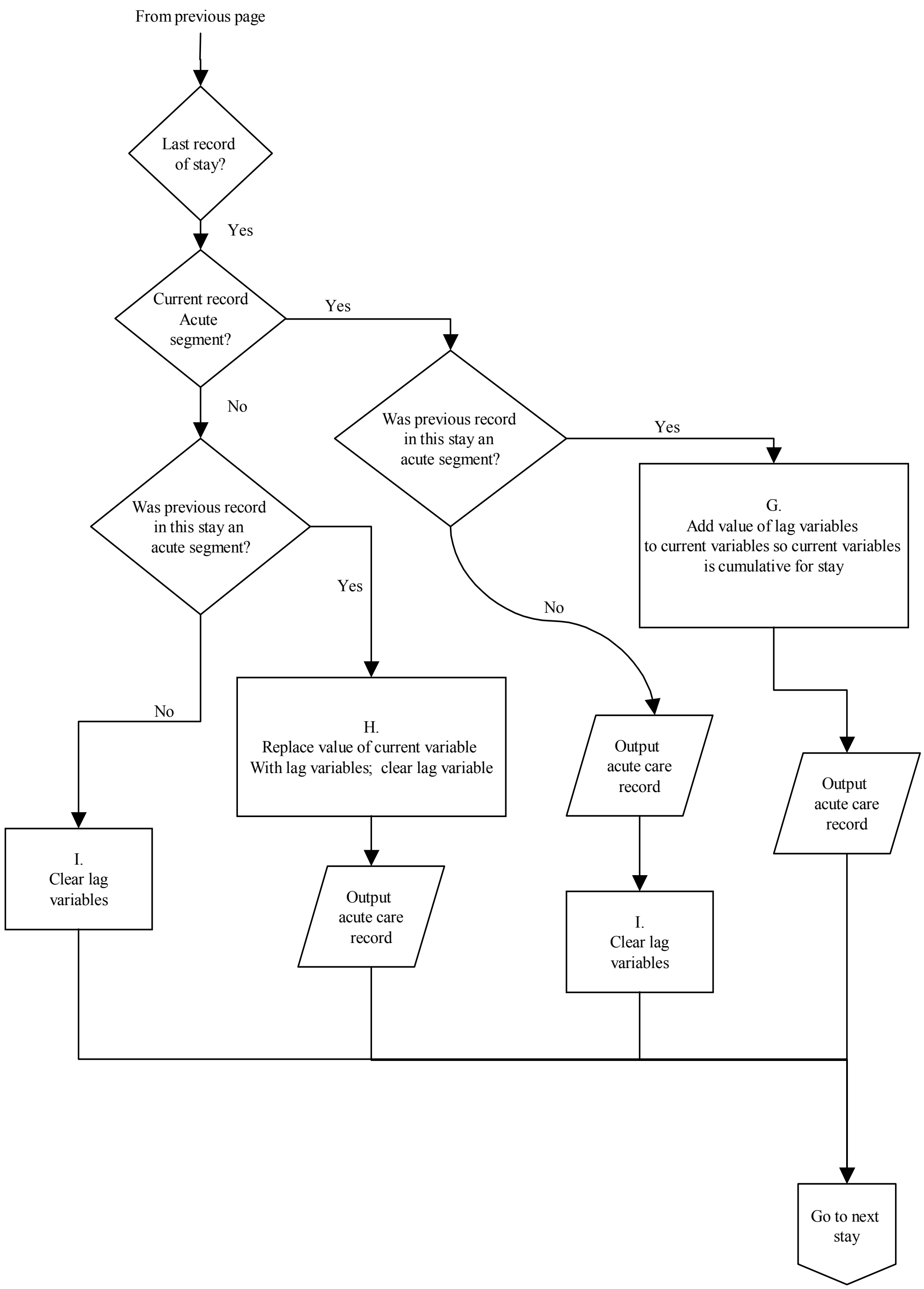
We then sorted the data by scrambled social security number (SCRSSN), medical center (STA3N), bedsection in day (BSINDAY) and bedsection out day (BSOUTDAY). Acute medical-surgical bedsection stays that were contiguous in time were considered to be part of the same hospitalization. Transfers within acute medical-surgical bedsections, such as from surgery to medicine, were aggregated into a single record. We adopted the rule that if a patient was transferred from an acute medical-surgical bedsection to another acute medical-surgical bedsection that this would be considered part of the same stay. Similarly, if a person was transferred from an acute medical-surgical bedsection to a non-medical/surgical bedsection, we ruled that the acute medical-surgical stay had ended. Transfers from an acute medical-surgical bedsection to a non-medical/surgical bedsection and back to an acute medical-surgical bedsection yielded one non-medical/surgical and two acute medical-surgical stays.

Figure 7.1 shows the flow diagram for how we compiled the acute medical-surgical hospital stays. The program starts with the PTF bedsection data and cycles through the records,

Figure 7.1: Accumulating contiguous acute medical-surgical bedsection stays



Accumulating contiguous acute bedsection stays (cont.)



accumulating contiguous acute medical-surgical bedsection stays. The program also performs a number of other important functions, such as recalculating length of stay, identifying the highest DRG weight from multiple bedsections (see section 7.2), and calculating number of days spent in intensive care (ICU). The program produces two discharge files, one for acute medical-surgical care and one for non-medical/surgical care. Appendix D has the SAS code for accumulating the stays.

7.2 Selecting the DRG and the relative value associated with a DRG

VA assigns a DRG to each bedsection segment of the hospital stay, and another DRG to the PTF main file, representing the DRG for the entire stay. The DRG is based on the principal diagnosis, the condition that is responsible for the patients' admission to the hospital.[6] The Health Care Financing Administration has developed a set of weights based on the DRG (DRG weights). These DRG weights are used to pay hospitals for Medicare patients.

We decided to use the DRG weights for our relative weights in the cost function. DRG weights are not part of the VA databases and need to be obtained from CMS and added to the VA files. Given that we had 1996 Medicare data, we merged the 1996 DRG weights from CMS with the PTF bedsection file. Then while we were making the acute medical-surgical VA hospital discharge file, the highest DRG weight across all bedsections was maintained. The rationale for this is that a private hospital would follow the same logic to maximize reimbursement.

We considered but did not use other relative value systems. We decided that the weights developed by states to pay for Medicaid and other patients are likely to reflect the patterns of practice in a specific state and that it would not be appropriate to apply them to the VA's national system of hospitals. Some relative value systems, such as the Severity of Illness Index, may provide some additional measure of relative cost (Averill et al., 1992), but they are not feasible for us to implement as they require data that are not available in VA utilization data at Austin. Patient Management Categories and Disease Staging are case-mix methods that can be applied to standard datasets, but they have been found to explain only 1-2% more variation than DRGs used alone (Calore & Iezzoni, 1987).

NEW For the FY02 cost estimates we used the 2002 DRG weight file from the Centers for Medicare and Medicaid Services (formerly known as the Health Care Financing Agency).

7.3 Length of stay

Length of stay is reported in the PTF bedsection file. But we had to recalculate length of stay according to our definition of acute medical-surgical stay (see section 7.1). Consequently, length of stay represents all days the patient spent in contiguous acute medical-surgical care bedsections during the stay.

---

[6] Prior to October 1, 1994, VA used the primary diagnosis to define DRGs. The primary diagnosis is the most important condition treated in the stay (as opposed to the principal diagnosis, which is the diagnosis responsible for the patient's admission to the hospital). VA DRGs from stays that ended prior to this date are thus not strictly comparable to non-VA DRGs from that time period, which have always used principal diagnosis.

7.4 Building the cost function

In past years we used an econometric method of estimating VA acute medical-surgical care costs (Barnett, 1997). Starting with FY98, we developed a cost function for estimating the cost of acute medical-surgical care. The cost-function is based on non-VA data, where the hospital stay as the unit of analysis. Using the stay (rather than the average stay) as the unit of analysis provides much more variation, including observations with high DRG weights and long lengths of stay. The cost function approach allowed us to construct a more complex model that better simulates the cost of stays with characteristics that are very different from the mean.

While the mechanics of the cost function are complicated, the intuition is relatively straightforward. We built a statistical model with a hospital discharge dataset. This regression model had cost adjusted charges on the left-hand side. On the right-hand side, we included variables such as length of stay, DRG weight, whether the patient died in the hospital, age, gender, and so forth. We saved the parameters from the regression model (i.e., the beta coefficients). This vector of coefficients was used to estimate costs in the VA data. It is important to note that the only way this approach can work is for both datasets to have the exact same right-hand side variables.

7.4.1 Data

We chose to use Medicare data for the cost function. Medicare data have some limitations, namely that Medicare does do not cover non-disabled individuals under age 65. For this reason, we carefully compared Medicare data to the Health Care Cost and Utilization Project (HCUP) data.

To provide some background on these datasets, the Medicare data were a subset of the 1996 Medpar file. The Medpar file was constructed by researchers at the Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC). They established a cohort of all veterans who were users of either inpatient or outpatients VA services between 1992 and 1994 and who had their 65th birthday in 1994. This cohort was then linked to Medicare denominator file to obtain Medicare enrollment. The file that we received represented 372,046 stays from hospitals in the continental US.

The HCUP data represents discharges from all types of hospitals in 22 states. Detailed information on the HCUP dataset is available on-line from www.ahrq.gov.

The primary question is, can we use the Medicare data to build a model that can estimate costs for younger veterans? Recall that Medicare data do not include non-disabled individuals under age 65. We answered this question by building a cost function with Medicare data. The function was then used to estimate the cost of stays in the HCUP sample. We then compared the estimated Medicare costs to the costs reported in the HCUP. This comparison was made for adults over 65 as well as adults under age 65. The remainder of this section describes this comparison.

First we selected a 40% random sample of non-ESRD Medicare claims in the MAVERIC cohort (125,457). With these claims, we estimated the following model:

CAC=a+b1died +b2sex +b3age+ b4npr+ b5npr2 +b6los + b7poslos + b8neglos + b9nlos2 +b10plos2 + b11nlos3 + b12drgwt +b13drgwt2 +e

where

CAC is cost adjusted charges

npr is number of surgical procedures
npr2 is number of surgical procedures squared
los is DRG specific length of stay
poslos is (average los-los) if average los > los
neglos is (average los-los) if average los < los
nlos2 nlos3 are square and cubic terms of neglos
plos2 is squared term of poslos
drgwt is CMS drgwt
drgwt2 is drgwt squared

The parameters from this model were saved and then used them to impute estimated costs for HCUP. We tried alternative model specifications, including the log transform of cost adjusted charges and excluding people with end stage renal disease (ESRD). In all of these alternative specifications, the parameters for the older people were remarkably similar to the parameters for the younger populations. We concluded that we could use the Medicare data to estimate the costs of younger hospitalized patients. The main advantage to this approach is that the Medicare data identify the number of days spent in intensive care. Because intensive care is resource intensive and costly, being able to estimate this parameter was a key advantage.

NEW For the FY01 & FY02 cost estimates, we received a 1999 Medpar file from the Massachusetts Veterans Epidemiology Research and Information Center. This is the file we used for estimating costs.

7.4.2 Cost adjusted charges

Utilization databases, like the Healthcare Cost and Utilization Project (HCUP) or Medicare, report charges incurred in a hospital. Yet, it is generally known that health charges usually exceed the cost of providing care. However, the degree to which charges exceed costs is not completely random. Hospitals and medical centers are somewhat idiosyncratic in how they generate bills.

Hence, we want to adjust the charges for two reasons: (1) to deflate charges so that they more closely reflect costs, and (2) to remove hospital specific idiosyncracies. The ratio of costs to charges (RCC), described in detail below, is one way of making this adjustment.

Adjusting charges with the RCC leverages information that every hospital annually reports to Medicare in the Medicare Cost Report. The Medicare Cost Report is a very large report that hospitals are required to complete if they want to receive Medicare reimbursements.

In the Medicare Cost Report, there are variables for each hospital's total charges and total costs. In the most recent Medicare Cost Report (PPS version 13), the field for charges is 2135 and the field for costs is 2138. We extracted these fields along with the hospital's Medicare identification number (PPS number). The quotient (i.e., the result of dividing costs by charges) was the ratio of costs to charges (RCC). The RCC usually ranges between 0.5 and 1.0. To actually adjust charges, the RCCs were linked to the Medicare dataset with the PPS number. The charge data were then adjusted by the RCC.

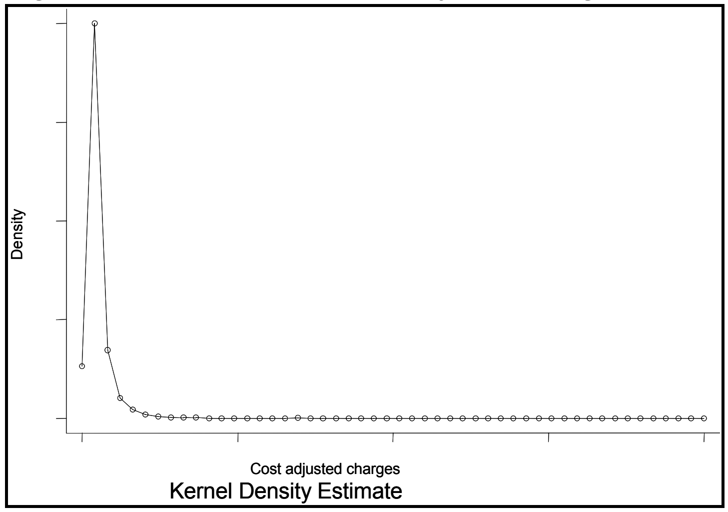
For example, if we want to use the RCC to adjust charges in a dataset, such as the HCUP dataset, we must first crosswalk the RCC dataset to the HCUP dataset. This can be a complicated process, especially for crosswalking the HCUP to Medicare (for details, see http://www.herc.research.med.va.gov/FAQ.htm). Once we crosswalk the files, we then multiply charges by the RCC. Recall that the RCC is a hospital-specific adjustment. In other words, within any given hospital the RCC will be constant.

For FY01 & FY02, we obtained the 1999 Medicare Cost Report (PPS16). PPS16 has different variables than the PPS13. They provide department level costs and charges. We used this to create a facility cost to charge ratio.

7.4.3 The dependent variable

We used cost adjusted charges as our dependent variable when we built the cost function. However, as shown in Figure 7.2, the cost adjusted charges from the Medicare data are not normally distributed.

Figure 7.2: Distribution of cost adjusted charges



Because of the skewness, we tried transforming the cost adjusted charges. While the log transformation helped reduce the appearance of skewness, the non-logged function consistently performed better than models with logged cost adjusted charges. Using logs presents additional hurdles because the estimated costs need to be transformed back to the original metric (dollars), adjusting for retransformation bias. The usual adjustment for retransformation bias is the smearing estimator (Duan, Manning, Morris, & Newhouse, 1983). While relatively simple to implement, this adds another layer of complexity to the entire process.

7.4.4 Length of stay

There are different ways to include length of stay in a cost function. The most obvious way is to include it without making any transformations, such that length of stay is a positive integer. Variations on this approach were also considered, such as a set of dummy variables representing different lengths of stay.

A second method for including length of stay involves comparing the patient's length to the average length of stay for all patients with that DRG. This second approach requires knowing the average length of stay for each DRG. This information is conveniently provided by CMS with the DRG weight file. We found slight advantages to the second approach as the transformation turned the length of stay from a positive integer into a continuous scale. Having a continuous scale provides slightly more ability to discriminate costs based on deviations in length of stay.

We used the second approach. In addition, we relaxed the constraints of our earlier estimates, allowing the cost of marginal days of stay to vary, depending on the length of stay.

Note that we examined only those records of patients discharged during the fiscal year under study. We included days of stay in acute medical-surgical bedsections, even if they occurred in previous fiscal years, and excluded data from stays that were not complete by the end of the fiscal year. This is distinct from the rest of our method, which considered only the days of stay that occurred during the fiscal year under study. We also calculated the length of stay in ICU bedsections. For each acute medical-surgical hospital stay, we found the number of days spent in the medical and surgical ICU bedsections.

7.4.5 Individual DRG intercepts or DRG weights

We found little marginal value in including dummy variables for each DRG. When we included DRG weight (squared and cubic terms), the gain in R^2 was less than 1%. Given the additional complexity in estimating this model, we decide to not use it. Instead, we decided to use DRG weight in our cost function along with the DRG weight squared and cubed. In the final model, we also interacted the Medicine Major Diagnostic Category (MDC) and Surgery MDC with length of stay.

7.4.6 Final model

The final cost function model based on a 50% sample of the Medicare data is shown in Table 7.1. The variable definitions follow.

7.4.7 Outliers

Outliers can have undue leverage on a regression model. After we ran the model, we found that the model fit the data reasonably well. However, the fit was based primarily on the high cost users. The model did not fit as well for low-cost users, due in part to heteroskedasticity.

One solution involves removing or "trimming" outliers. We tried this and retested the model fit. Our methods and findings are below. We first identified outliers by using the Medicare outlier designation (n=1880). This did not help the fit of the model with low-cost cases because the outlier designation typically identifies the expensive cases.

Then we empirically identified outliers by generating Cooks' distance. Cooks' distance is the leverage of case *I* on the OLS regression coefficients (*$hat*). It can be thought of as an F test comparing the beta coefficients with and without observation *I (i.e., $hat* to *$hat*$_{-I}$). Large values for Cook's distance suggest that the case has a lot of leverage.

We trimmed outliers in our regression models using three exclusion criteria:[7]

1) Cooks distance >0.001 (excluded 968 observations, ~0.8%)
2) Cooks distance >0.0001 (excluded 2,101 observations, ~1.7%)
3) Cooks distance >0.00001 (excluded 8,431 observations, ~6.6%)

We found that we could estimate better fitting models if some outliers were excluded. This gain was mainly within the lowest quartile of costs. Table 7.2 presents correlation coefficients between actual cost adjusted charges (CAC) and estimated cost adjusted charges. Note, however, that not always did removing more outliers lead to a better fitting model. In quartile 1, only model #3 yielded higher correlations.

[7] We also compared logged CAC models. In every case, the log models fit significantly worse and yielded much larger differences between estimated costs and actual costs.

Table 7.1: Full model based on 50% random sample of Medicare data (FY98-00)

Source	SS	df	MS
Model	3.8009e+13	27	1.4078e+12
Residual	1.3554e+13	321555	42152405.8
Total	5.1564e+13	321582	160343662

Number of obs	= 321583
F(27,321555)	=33396.73
Prob > F	= 0.0000
R-squared	= 0.7371
Adj R-squared	= 0.7371
Root MSE	= 6492.5

cac	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
died	2671.211	57.21167	46.690	0.000	2559.077	2783.344
sex	32.90875	61.21531	0.538	0.591	-87.0715	152.889
age	-34.22324	1.851834	-18.481	0.000	-37.85278	-30.5937
ndx	619.0444	81.09738	7.633	0.000	460.0959	777.993
ndx2	-146.7017	16.61743	-8.828	0.000	-179.2714	-114.1321
ndx3	10.97541	1.022981	10.729	0.000	8.970401	12.98043
los	104.255	9.083375	11.478	0.000	86.45187	122.0582
poslos	670.9503	10.10664	66.387	0.000	651.1415	690.759
neglos	182.4991	29.68224	6.148	0.000	124.3228	240.6755
nlos2	-109.8903	7.980714	-13.769	0.000	-125.5323	-94.24832
plos2	-.7170458	.021736	-32.989	0.000	-.7596478	-.6744437
nlos3	-4.587643	.5484962	-8.364	0.000	-5.66268	-3.512606
plos3	3.32e-06	.0000198	0.168	0.867	-.0000354	.000042
drgwt	4860.036	63.69243	76.305	0.000	4735.201	4984.871
drgwt2	-255.1638	11.0401	-23.112	0.000	-276.8021	-233.5255
drgwt3	12.97284	.5057919	25.649	0.000	11.98151	13.96418
surg	1069.883	78.21631	13.679	0.000	916.581	1223.184
surlos	-42.31538	11.16155	-3.791	0.000	-64.19169	-20.43906
pl_sur	421.5315	15.61753	26.991	0.000	390.9216	452.1415
nl_sur	328.304	36.252	9.056	0.000	257.2511	399.3569
pl_sur2	-1.384451	.1793446	-7.720	0.000	-1.735961	-1.03294
pl_sur3	.001167	.0006719	1.737	0.082	-.00015	.002484
nl_sur2	47.49814	8.419396	5.642	0.000	30.99636	63.99991
nl_sur3	3.636805	.55208	6.587	0.000	2.554745	4.718866
icudays	593.0367	7.165874	82.758	0.000	578.9918	607.0816
icudays2	10.27421	.2713893	37.858	0.000	9.742298	10.80613
icudays3	-.0325464	.0017843	-18.240	0.000	-.0360436	-.0290492
_cons	413.7664	181.3739	2.281	0.023	58.27884	769.254

The variables in the model are:

- died: died in hospital
- sex: 0=male, 1= female
- age: age in years
- ndx: number of diagnoses
- ndx2: number of diagnoses squared
- ndx3: number of diagnoses cubed
- los: length of stay in days
- poslos: positive deviation from DRG specific average LOS
- neglos: negative deviation from DRG specific average LOS
- nlos2: neglos squared
- plos2: poslos squared
- nlos3: neglos cubed
- plos3: poslos cubed
- drgwt: CMS 1996 DRGweight
- drgwt2: DRGweight squared
- drgwt3: DRGweight cubed
- surg: surgical MDC
- surlos: surgical MDC* LOS
- pl_sur: surgical MDC* poslos
- nl_sur: surgical MDC* neglos
- pl_sur2: surgical MDC* plos2
- pl_sur3: surgical MDC* plos3
- nl_sur2: surgical MDC* nlos2
- nl_sur3:surgical MDC* nlos3
- icudays: days in ICU
- icudays2: icudays squared
- icudays3: icudays cubed
- _cons: constant

We decided not to remove outliers because we realized any decision about which outliers should be removed would be arbitrary and would affect the model's fit. The full model fits almost as well (and better in some instances), therefore we saw little rationale for removing outliers.

Table 7.2 also shows how well the model predicts costs with the other 50% of the data (out of sample). In many cases, the out-of-sample predicted costs are quite close to the actual Medicare costs. As is shown in Table 7.1, the overall R^2 of the model is approximately 0.74.

Table 7.2: Correlations between estimated costs and actual costs for the full model and for three outlier restricted models

	Actual costs							
	Quartile 1: <$2605		Quartile 2: $2605<cac<$4484		Quartile 3: $4484<cac<$8472		Quartile 4: >$8472	
	In sample	Out of sample	In sample	Out of sample	In sample	Out of sample	In sample	Out of sample
Sample size	38304	38144	39167	38594	39939	40801	43348	43286
Model with all cases				**correlation coefficients**				
estimated costs	0.126	0.190	0.301	0.291	0.389	0.357	0.814	0.808
Restricted models								
(1)	0.057	0.204	0.309	0.005	0.396	0.250	0.641	0.699
(2)	0.071	0.209	0.313	0.011	0.398	0.279	0.718	0.749
(3)	0.185	0.202	0.313	0.305	0.393	0.392	0.769	0.775
Model estimated with log(CAC)	0.083	0.109	0.303	0.290	0.390	0.381	0.389	0.106

Notes: (1) cost function was estimated excluding cases with a cooks' distance >.001 (least restrictive)
(2) cost function was estimated excluding cases with a cooks' distance >.0001 (more restrictive)
(3) cost function was estimated excluding cases with a cooks' distance >.00001 (most restrictive)

7.5 Observation days

Beginning in 1997, VA created 7 new codes for observation bedsections to report inpatient care provided in observation units. Most stays involving these codes are recorded in the observation PTF files, which is a new set of files in the PTF. These stays, even if there are associated with an inpatient record in the Acute PTF file, are kept in a separate observation bed file at Austin. The structure of the observation files mirror the PTF inpatient files. We found that many stays reported in this file precede or follow stays in the acute medical-surgical PTF file. When calculating length of stay, some analysts will want to regard these observation days as part of acute medical-surgical stays.[8]

[8] Nearly 73,000 days of stay were assigned to observation bed sections in FY99 (out of 13.5 million days in VA hospitals). Most of the observation stays were one day long, but this was not always the case. Most observation days were in medicine, surgery, and psychiatry observation bedsections. We recently examined the FY99 data and found that 19,428 (26%) of

For the cost of observation bed stays, we decided that for FY98 onward we would cost each day at the marginal cost of an additional day (i.e., $684; see section 7.6). This method may underestimate the cost of stand-alone observation stays. Alternatively, it may overestimate the cost of an observation stay that preceded a hospitalization. We hope to develop and test new methods for costing observation bed stays in the future.

7.6 Negative or implausible costs

After estimating VA costs with the cost function (see Table 7.1), we found that the function had imputed negative costs for 2,974 of the 541,567 (0.6%) acute medical-surgical hospitalizations. This is because the cost function was not constrained to predict non-negative estimates. Therefore, rare combinations of right-hand-side variables can lead to negative predictions. These 2,974 records were assigned the cost of a marginal day of stay ($684.75).

The cost of a marginal day of stay was calculated in a simulation with the 1996 Medicare data. Adjusting for all other covariates in a linear regression, we identified the cost for an additional day of stay. Holding all other factors at their mean, if a person stayed an additional day, they had an additional $684.75 of cost adjusted charges.

While some stays were not assigned negative costs, they were given very low costs. For instance 42 hospital stays had positive costs less than $5. We decided that any stay with a cost less than $684.75 was implausibly low and an artifact of the cost function. By setting this rule, it effectively set a floor on the estimated cost per stay. A total of 9,632 (2%) cases had non-negative costs less than $684.75. These cases were all given $684.75 per day (86% had a length of stay of one day). In the future, we will explore other methods for determining the cost of these cases, including setting constraints on the cost function.

7.7 Reconciling to the CDR

The cost function is based on non-VA relative value weights and non-VA cost adjusted charges. The estimated costs must be reconciled to the Cost Distribution Report to reflect VA costs. Reconciliation can happen at many levels including the department, medical center, and nationwide. We chose to reconcile the estimated costs to the medical center and nationwide; we decided not to reconcile the estimated costs to the department. Given that the CDR and PTF are not reconciled against each other, our concern was that there would be too much variability in department-level costing.

Reconciling the costs to the medical center results in “local” cost estimates, while reconciling the costs for the entire VA results in “national” cost estimates. Therefore, this process results in the creation of 2 VA cost estimates: a local cost estimate (costl) and a national cost estimate (costn).

The logic behind reconciling the costs is straightforward. For the local cost estimate we sum together the estimated costs for a medical center and divide this amount by the total acute medical-surgical care CDR costs (acute medicine and surgery) for the medical center. The quotient of this division is a scaling factor. By multiplying the estimated cost by this scaling

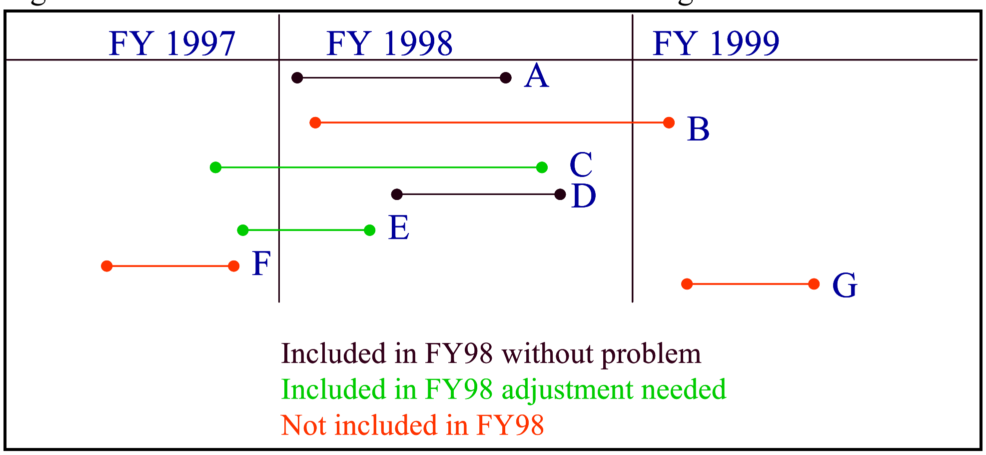
---

the observation stays immediately preceded a stay reported the PTF bedsection files. Another 319 observations stays followed stays in the bedsection file. (Our analysis was limited to PTF bedsection file. It is also possible that observation stays precede or follow stays reported in the PTF extended care file.)

factor, we ensure that the sum of the estimated costs is equivalent to the CDR costs.

Unfortunately, the reconciliation is easier said than done. Recall that the CDR reports costs for the fiscal year while the acute medical-surgical hospitalization data represent discharges. For FY98 data, some stays that ended in FY98 started before FY98. At the same time, there were people hospitalized in FY98 who were still in the hospital at the end of the fiscal year and are not reported in the FY98 PTF data. To illustrate this point, Figure 7.3 shows the hospitalization that cross the fiscal years. Cases B, C, and E all cross the fiscal years. It is not correct to assume that the cases crossing from FY97 to FY98 are equivalent in number to those cases crossing from FY98 to FY99. Due to the declining trend in inpatient hospitalization, C and E are more common than B.

Figure 7.3: Difference between FY view and discharge view



Note: A & D are in the med/surg file and need no adjustment
C & E are in the med/surg file and need adjustment
B, G, and F are not in the med/surg file

If no adjustment were made for this fact, then we would overestimate the number of hospitalizations, and thereby underestimate the cost of care per hospitalization. Our correction for this was to adjust the cases discharged in the fiscal year that started before the fiscal year. For example, for FY98, the adjustment factor was found by comparing the FY98 Census to the FY 97 Census. See chapter 3.1.3 for a discussion of the census files. This comparison showed that the acute medical surgical- stays crossing the 1998/99 fiscal year were 93% as frequent as the cases crossing the 1997/98 fiscal year. Consequently, we used 0.93 as our adjustment factor and multiplied the estimated costs for cases crossing the 1997/98 fiscal year (i.e., points C and E in Figure 7.3) by this factor. The FY98/99 ratio was 0.9821, the FY00/99 ratio was 0.9290, and the FY01/00 ratio was 1.0442.

After adjusting the discharge data so that it better represented the FY costs in the CDR, we reconciled the estimated costs. The national scaling factor was 1.2627, 1.290, and 1.405 in FY98, FY99 and FY00, respectively. In other words, we multiplied every estimated cost by this scaling factor to obtain the national VA cost. This ensures that if every acute medical-surgical hospitalization discharged in the fiscal year were summed together that the total would equal the CDR costs. The local scaling factor was different for each medical center. The average local scaling factor was 1.262, 1.285, 1.400, 1.42 in FY98, FY99, FY00, and FY01 respectively.

7.8 Stability of the cost function over time

The cost function for FY98-FY00 was built using 1996 Medicare data. In FY01, we used 1999 Medicare data. One question is whether the cost-function is robust to the input data that are being used. To answer this question, we loaded 1994 and 1995 Medpar data that was similar to the 1996 Medpar data. We then ran the identical cost function on all three datasets. The model coefficients from the three datasets were compared. Finally, using the regression model for each year of data, I predicted costs in 1996, using the Medpar 1996 as the criterion. We compared the estimated costs to see if differences would have occurred had they been estimated with 1994 or 1995 Medpar data.

The regression coefficients for all three models were extremely similar (Table 7.3). The predicted costs from the three models were also highly correlated (>0.99; Table 7.4). The results suggest that the cost function is highly robust to the year from which the Medpar data are used.

Table 7.3: Stability of regression coefficients with 1994, 1995 and 1996 Medpar data

	1994		1995		1996	
	coeff	t-stat	coeff	t-stat	coeff	t-stat
died	2837.70	42.410	2803.32	42.650	2671.21	46.690
sex	-41.01	-0.560	-28.73	-0.400	32.91	0.540
age	-42.29	-18.590	-44.42	-19.720	-34.22	-18.480
ndx	250.36	2.740	433.47	4.710	619.04	7.630
ndx2	-80.63	-4.190	-117.71	-6.180	-146.70	-8.830
ndx3	7.44	6.150	9.60	8.120	10.98	10.730
los	50.63	4.660	52.01	4.890	104.26	11.480
poslos	656.08	54.620	666.76	54.250	670.95	66.390
neglos	272.94	9.400	338.59	11.140	182.50	6.150
nlos2	-72.45	-11.940	-71.91	-10.220	-109.89	-13.770
plos2	-1.31	-54.080	-0.62	-10.450	-0.72	-32.990
nlos3	-1.41	-4.830	-1.85	-4.490	-4.59	-8.360
plos3	0.00	30.680	0.00	2.900	0.00	0.170
drgwt	4477.58	58.500	5149.17	69.610	4860.04	76.300
drgwt2	-161.85	-12.100	-325.22	-25.390	-255.16	-23.110
drgwt3	8.02	13.030	16.71	28.480	12.97	25.650
surg	470.37	5.280	526.47	5.890	1069.88	13.680
surlos	-48.96	-3.770	-23.43	-1.810	-42.32	-3.790
pl_sur	416.50	26.280	379.25	22.240	421.53	26.990
nl_sur	222.54	5.670	152.01	3.850	328.30	9.060
pl_sur2	-1.21	-24.520	-0.95	-8.300	-1.38	-7.720
pl_sur3	0.00	18.310	0.00	-1.250	0.00	1.740
nl_sur2	18.26	2.590	3.07	0.390	47.50	5.640
nl_sur3	0.58	1.900	0.72	1.710	3.64	6.590
icudays	395.04	47.070	553.12	67.840	593.04	82.760
icudays2	18.93	58.260	9.29	31.130	10.27	37.860
icudays3	-0.08	-37.720	-0.02	-11.440	-0.03	-18.240
cons	1819.08	8.640	1416.06	6.650	413.77	2.280

Table 7.4: Pair wise Correlations in predicted costs compared to 1996 costs adjusted charges

	cost94	cost95	cost96
cost94	1.000		
cost95	0.993	1.000	
cost96	0.997	0.996	1.000
CAC 1996	0.856	0.855	0.859

Note: CAC is cost adjusted charges

Chapter summary

- To estimate the costs of acute medical-surgical care for FY98-FY01, we developed a cost-function from Medicare Medpar data restricted to Veteran users.

- HERC developed a VA acute medical-surgical dataset using the PTF bedsection file. Contiguous acute medical-surgical bedsection stays were aggregated into a single record. This program also recalculates LOS, ICU days, and keeps the highest DRG weight for all acute medical-surgical bedsection stays.

- In building the cost function, we compared the HCUP dataset to a veteran-restricted Medicare dataset. The Medicare dataset was able to predict the costs of younger people in the HCUP dataset and it identifies ICU days, which are a useful indicator of resource use. Therefore, we used the veteran-restricted Medicare dataset.

- Medicare reports charges. We adjusted the reported charges with a hospital-specific ratio of costs to charges. This effectively deflates the reported charges and removes hospital-specific billing differences.

- Length of stay was entered into the model as the deviation from the expected length of stay for that DRG.

- After comparing alternative models, we decided to use DRG weight as the measure of relative weight, rather than allow each DRG to have its own intercept.

- The 1996 Medpar model had an R^2 of 0.7371. The 1999 Medpar model was 0.7539.

- We explored whether to trim influential outliers. This affected the model's fit, and not always positively. Because the cut-off for selecting the outliers was arbitrary, we included all cases.

- For each observation day, we costed it at the marginal cost per day, which we estimated at $684.75.

- The cost function yielded some negative and implausible costs. We set $684.75 (the marginal cost of a day), as the minimum cost possible.

- We reconciled the estimated costs to the CDR for the medical center and the nation. This yielded a local cost estimate (costl) and a national cost estimate (costn).

Chapter 8. User's Guide

This chapter discusses how to use HERC's average cost dataset. The chapter is broken into four sections: 1) a brief summary of the methods, 2) assumptions underlying the dataset, 3) how to correctly use the dataset, and 4) when not to use the dataset. We strongly feel that every user of these data should be knowledgeable in these areas.

In this chapter, we do not spend much time discussing the potential uses of the average cost data. We expect that these data will be extremely useful for many, if not the majority of, VA-research projects. In fact, these data have already been used in trend and econometric analyses. In addition, a number of clinical trials are planning on using these data. While we hope that these data will be useful, we do not expect that these data will be appropriate for every study. For this reason, later in this chapter we discuss limitations with these data and instances for which these data are not appropriate.

8.1 Summary of methods

8.1.1 Categories of inpatient care

Starting in FY98, we categorized inpatient care into eleven categories: 0) acute medicine, 1) rehabilitation, 2) blind rehabilitation, 3) spinal cord injury rehabilitation, 4) surgery, 5) psychiatry, 6) substance abuse care, 7) intermediate medicine, 8) domiciliary, 9) nursing home care, and 10) psychosocial residential rehabilitation programs (PRRTP). These categories are defined by bedsection (see Table 4.3). While PRRTP care is defined by bedsection, it is only available at approved medical centers. If a non-approved medical center had dollars or days in PRRTP bedsections, these were allocated back into psychiatry and substance abuse care, respectively.

8.1.2 Acute medical-surgical care

Of the eleven categories of care, acute medicine and surgery comprise the acute medical-surgical care. For patients receiving this type of care, we estimated costs using a cost-function from Medicare MedPar data restricted to Veteran users (see Chapter 7). To do this, we developed a VA acute medical-surgical dataset using the PTF bedsection file. Contiguous acute medical-surgical bedsection stays were aggregated into a single record.

In building the cost function, we used a veteran-restricted Medicare (MedPAR) dataset. We adjusted the reported Medicare charges with a hospital-specific ratio of costs to charges. In the cost function, length of stay was entered into the model as the deviation from the expected length of stay for that DRG. We also used DRG weight as the measure of relative weight, rather than allow each DRG to have its own intercept.

For each observation day in an acute medicine or surgical bedsection, we costed it at the marginal cost per day, which we estimated at $684.75. The cost function yielded some negative and implausible costs. We set $684.75 (the marginal cost of a day), as the minimum cost possible.

Lastly, we reconciled the estimated costs to the CDR for the medical center and the nation. This yielded a local cost estimate (costl) and a national cost estimate (costn).

8.1.3 Nursing home care

For FY98-FY00, nursing home costs reflect case-mix. Using the Resource Utilization Groups (RUGs) that are collected biannually on nursing home patients, we imputed the daily cost per RUG unit. To obtain the patient's cost per stay, we multiplied each patient's rug score by the per rug cost times the length of stay. The methods for this are discussed in detail in Chapter 6.

Nursing home costs for FY00 were based on an unadjusted per diem. In FY01, VA started using the RUG/MDS data collection tool, rather than the RUG II score. The RUG III data are not yet available.

8.1.4 Non medical/surgical categories

All remaining cost categories were estimated as a daily rate. The total CDR costs were divided by the total units provided in the PTF bedsection file. The daily rate methods are described in detail in Chapter 5.

8.2 Assumptions in the average cost dataset

Throughout this document we have tried to identify assumptions underlying the creation of the acute medical-surgical and non medical/surgical datasets. Both datasets reconcile to the CDR at the level of the medical center and the nation. Costs excluded from the CDR are also not included in our estimates. These include, importantly, the cost of financing capital expenditures and malpractice costs. Our average cost estimates do include indirect costs and physician costs. Table 8.1 shows the included and excluded costs.

Table 8.1: Included and excluded costs

Type	Notes
Excluded	
Capital financing costs	Not included, but this may be noteworthy (5%).
Malpractice expenses	Not included.
Contract provider costs	Excluded are contract services because these costs are not accurately associated with units of care
Community nursing home costs	We excluded cases that were in bedsection 80 with Statyp 42.
Headquarters costs	Excluded are the costs associated with VA headquarters
Prosthetics	Inpatient prosthetics billed separately are not included in the CDR accounts
Included	
Costs for physician services	These costs are included in the CDR. For every stay, physician costs are proportionate to the hospital costs.
Research & education	Included to the extent supported by the VA medical care appropriation.
Indirect costs	We assigned indirect costs to each CDA in proportion to its share of the total direct costs of its group of CDAs.

8.2.1 Data used in the cost function

The average cost estimates for acute medical-surgical stays were based on a cost function that was constructed with Medicare data. The 1996 and 1999 Medicare data represented veteran users; excluded were cases in Hawaii, Alaska and cases related to labor and delivery. In using the Medicare data we assumed that the underlying accounting systems for non-VA hospitals could be used to impute estimates for the VA. These imputed estimates were then reconciled with the CDR. If you were to sum all of our cost estimates for a medical center in a given year, you will find that the local cost total is equivalent to the amount posted in the CDR.

8.2.2 The cost of observation stays

Observation stays are a relatively new type of service provided in the VA. There is no analogous type of service provided in the private sector. To estimate the cost of the observation bed stay, we estimated a marginal daily rate and multiplied this times the length of stay. Most people stay in the observation bed for one day; a few outliers stay longer and in these cases, the cost is equivalent to this rate times the length of stay. To calculate the daily rate for observation bed stays, we developed a regression model using Medicare data. With the regression model, we simulated the marginal cost at the mean of data. We then predicted the cost if the person stayed one day longer than the mean. The difference between these two estimates was \$684.75. We used this as the daily rate for the observation bed stays.

8.2.3 Costs for high and low-cost procedures

The cost function used to estimate acute medical-surgical costs was presented in chapter 7. As was mentioned in that section, the model does a better job estimating high cost stays. The accuracy of the average cost estimate is better with high-cost cases than with low-cost cases. If you are assessing cases that typically have very low costs, then the average cost provided in the HERC dataset may be inappropriate.

8.2.4 Implicit trimming of outliers

A byproduct of using the cost function is that it removes outliers. Recall that the cost function is a linear regression model. When we calculated the cost for the VA we used the regression model to estimate costs based on averages. If you are interested in high or low-cost outliers, then the HERC dataset may be inappropriate for your use.

8.2.5 Model estimates and negative costs

Another byproduct of using a cost function is that after we imputed the VA costs we had some cases with negative or implausibly low costs. Clearly, a stay cannot have a negative cost. Therefore, we decided that we would set a floor. Any choice of a floor is somewhat arbitrary, but we chose the floor to be \$684.75. Recall that \$684.75 is the average cost of an additional day of stay (see chapter 7). A total of 12,731 cases had an estimated cost of less than \$684.75. For all these cases, we assigned them a cost of \$684.75. This cost was their total cost, NOT a daily rate. Of these cases, 83.5% (10,636) had only one day of stay. Another 14% and 2% had a stay of two and three days, respectively. The remainder (101 cases) had up to 8 days of stay; however, there were three outliers who had more than 1000 days of stay. Clearly a cost of \$684.75 is inappropriate for someone who stayed 1150 days in the hospital, but we did not make adjustments for these three cases. When you use these costs, compare the length of stay to the cost. Make sure that these three cases are not in your data. If they are, you probably want to exclude them or assign them a different cost.

8.2.6 VISN administrative costs

Each of the VISNs incurs administrative operating costs. We have included these costs under the assumption that they cover coordination expenses required for a large health provider. In the CDR, these costs are assigned to a single medical center within the VISN. From our perspective, these costs should be distributed to all medical centers in the VISN. We are looking into ways of distributing these costs, but for FY98-FY01, these costs remain where they were assigned. This may partly explain deviations in the local costs. This provides a reason for using national costs, but if your study requires local costs, then use them carefully.

8.3 Using the average cost dataset

At Austin, we have provided three datasets. These datasets are listed in Table 8.2 and described below.

Table 8.2 The three average cost datasets for FY98-FY00

Dataset	Includes	Excludes
dischgXX	• All persons admitted since FY98 and discharged in fiscal year. • Costs for acute medical-surgical are combined with non medical-surgical costs when bedsection stays within a discharge are contiguous.	• stays not completed by end of fiscal year • stays admitted before beginning of FY98 (10/1/97)
mdsrgXX	• All persons discharged from an acute medical-surgical bedsection in fiscal year	• Non medical-surgical bedsections • People who were still in the hospital at end of FY.
nmdsrgXX	• The cost of care provided in non medical-surgical bedsections during the fiscal year.	• The costs of care provided before the fiscal year are excluded.

8.3.1 Two important variables: source and flag

The datasets have a couple variables that may be of interest to users. First, as we compiled the datasets from multiple datasets, we kept track of where the data were from. The variable *source* indicates the case's source. The format for source is: 1=XB Census, 2=XB discharge, 3=PB Census, 4=PB discharge, 5=OBS discharge, 6=OBS Census.

Another important variable is the flag variable. This variable indicates when the local cost estimate (costl) is > 2 standard deviations above or below the national cost estimate. Flag is an indicator or dummy variable; use the costl with caution when the flag variable is one.

8.3.2 Discharge dataset

The discharge dataset was generated by combining the acute and nacute datasets. It represents a discharge dataset, such that it only has cases that were discharged. In addition, only people admitted since the beginning of FY98 are included in the discharge datasets. Patients that were admitted prior to FY98 are excluded

The discharge dataset includes additional variables that track cost subtotals, length of stay subtotals, DRG weight, and ICU days.

Subtotals are based on the following categories of care

0) Acute medicine/surgery,
1) Rehabilitation,
2) Blind rehabilitation,
3) Spinal cord injury rehabilitation,
4) Does not exist (this was surgery, but it has been combined with medicine to form category 0)
5) Psychiatry,
6) Substance abuse care,
7) Intermediate medicine,
8) Domiciliary,
9) Nursing home care, and
10) Psychosocial residential rehabilitation programs (PRRTP).

Variables

- COSTL_0 -COSTL_10: local cost estimates for categories of care 0 -10
- COSTN_0-COSTN_10: national cost estimates for categories of care 0 -10
- LOS_0 -LOS_10: length of stay estimates for categories of care 0 -10
- DRGWT: the diagnosis related group (DRG) weight. In cases where a stay was assigned more than one DRG weight, the corresponding observation only records the maximum DRG weight.
- ICUDAYS: number of days in an intensive care unit.

Due to this change, a single discharge record provides important subtotals. For example, if a researcher is interested in mental health costs, he/she can now identify the mental health costs for every inpatient encounter. This is particularly helpful for those patients who receive care in many different categories during a stay. Again, note that these changes only pertain to the inpatient discharge datasets.

8.3.3 Acute medical-surgical dataset

This dataset is best described as a discharge dataset for persons who were discharged or transferred from an acute medical-surgical bedsection in FY98-FY01. The key to understanding this dataset is that we aggregated the bedsection files to make a discharge file that is analogous to the MedPar dataset (see section 7.1).

The first step of the process involved identifying acute medical-surgical bedsections.[9] If, in a stay,[10] a person was in three acute medical-surgical bedsections, we combined these bedsections. Transfers within acute medical-surgical bedsections, such as from surgery to medicine, were aggregated into a single record. We adopted the rule that if a patient was transferred from an acute medical-surgical bedsection to another acute medical-surgical bedsection that this would be considered part of the same acute medical-surgical stay. Similarly, if a person was transferred from an acute medical-surgical bedsection to a non-medical/surgical bedsection, we ruled that the acute medical-surgical stay had ended. Transfers from an acute medical-surgical bedsection to a non-medical/surgical bedsection and back to an acute medical-surgical bedsection were treated as one non-medical/surgical and two acute medical-surgical stays.

You will want to link this file to the PTF bedsection files. But before you merge those files with this cost file, you will need to aggregate the bedsection file. We have provided the code for this in Appendix D. You can also contact HERC if you would like an electronic version of this SAS code.

8.3.4 Non medical-surgical dataset

This dataset contains costs for people who were in non medical-surgical bedsections. Only costs for stays during the fiscal year were included. If a person was admitted and discharged in FY98, then the total cost of their stay is in this FY98 dataset. However, if a person was admitted prior to FY98 (10/1/97), then only the costs for the portion of the stay during FY98 is reported in the dataset. One of the reasons for doing this is that there are some people in long-term care who have been there for 30+ years. It would be extremely difficult to identify the entire cost of these stays. For information on costs prior to FY98, see HERC working paper (Barnett, Chen, & Wagner, 2000).

[9] The medical-surgical bedsections in FY 98 were 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 14, 15, 16, 17, 18, 19, 24, 31, 34, 35, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 65, 75, 83. All others were considered were considered non medical-surgical.

[10] Stays were defined by five variables: scrssn, sta3n, admitday, adtime, disday.

Table 8.3 Using the three average cost datasets

Dataset	Sort and merge using	Merge data to
dischgXX	SCRSSN, ADMITDAY, DISDAY, and STA3N.	PTF main files (PM, XM and PMO)
mdsrgXX	SCRSSN, ADMITDAY, DISDAY, STA3N, and BSOUTDAY.	PTF bedsection files (PB, XB, PBO); BUT must first aggregate the bedsection file
nmdsrgXX	SCRSSN, ADMITDAY, DISDAY, STA3N, BSINDAY, and BSOUTDAY.	PTF bedsection files (PB, XB, PBO), and PTF census files.

8.4 When not to use the average cost dataset

8.4.1 Effects not detected in this cost estimate

It is not always appropriate to use these average cost data in your analysis. The average cost method assigns the same cost to all inpatient stays with the same demographic and discharge information. Stays that have the identical characteristics will have the same cost. If you are interested in assessing the cost consequences of a new procedure, then these data are likely to be inappropriate unless the cost of the procedure is entirely reflected by variables in the cost function (see page 36). If the procedure saves money, but it does not affect one of the variables in the cost function, such as DRG weight or length of stay, then these stays will all get the average cost.

For example, let us assume that we had a new procedure for transfusing blood during a heart transplant. We are interested in whether this new procedure saves money. First, let us assume that this intervention would not affect the patient's DRG. In this case, it is also likely that the intervention would not affect other variables in the cost function, such as length of stay. Therefore, the estimated cost of care for people who received this new procedure would be the same estimated cost of care for people receiving the usual therapy. This does not mean that there was not a cost difference from this new therapy. It only means that any differences were not reflected in the HERC average cost dataset.

8.4.2 Comparison of medical center efficiency

The economic definition of efficiency is to use fewer inputs to make the same level of output, or conversely to use the same number of inputs to make more output. These costs estimates are relative value weights based on Medicare patient discharge characteristics. The local cost estimate is generated by reconciling the relative value weights to the CDR. But, the relative value weights DO NOT capture differences in the quantity or price of the inputs. In addition, the CDR costs exclude the cost capital financing. In addition, we distribute other short-term fixed costs in proportion to the variable costs. Although these issues may not be critical for cost-effectiveness analysis, they are more problematic and potentially fatal for efficiency analysis.

8.4.3 Point estimates versus variance estimates

We believe the average cost method produces relatively accurate point estimates for the costs. However, a consequence of estimating costs with a cost function is that the variance of the estimated costs is biased downwards. The reason for this is that many factors that affect costs are not included in the cost function, and if the stays are identical on all observed factors then these cases receive the same estimated cost. In Table 8.4 we show the costs reported by Medicare (1996) for five DRGs. We also show the estimated costs from our cost function (estcost). As is clear from this table, the standard deviation is smaller in the estimated costs. Also, the minimum and maximum are attenuated toward the mean.

Table 8.4: The cost function's effect on the variation of the estimated costs

	Obs	Mean	Std. Dev	Min	Max
DRG14 Specific cerebrovascular disorders except TIA					
cost	10534	6829	7587	7	175346
estcost	10534	7377	7476	685	147135
DRG79 Respiratory infections & inflammations age >17 w cc					
cost	7767	7923	8445	16	213967
estcost	7767	8210	6423	685	198091
DRG88 Chronic obstructive pulmonary disease					
cost	15428	4786	5525	5	203877
estcost	15428	4535	4269	685	128695
DRG89 Simple pneumonia & pleurisy age >17 w cc					
cost	12905	5468	8863	8	662916
estcost	12905	5238	4675	685	160280
DRG127 Heart failure & shock					
cost	21463	4941	4979	10	109945
estcost	21463	5224	4479	685	190673

Note: cost is cost adjusted charges and estcost is the estimated cost adjusted charges.

If you are interested in evaluating the variation of these cost estimates, then use these costs carefully. If you use these cost estimates in a statistical model, most statistical tests will be biased toward the null. If you are trying to identify cost outliers (high or low), then you will almost certainly miss some.

8.5 Duplicates

Researchers who want to merge VA utilization data to our average cost estimates need to be aware that the PTF files have duplicates. There are duplicates within each file (e.g., PB discharge file) and between files (e.g., PB discharge file and XB discharge file). We have removed all duplicates in the average cost datasets before we calculated the costs. To prevent a one-to-many merge, you should delete duplicates from any Austin data that you are working with. The best way to handle this is to run the following command in SAS, which will remove any duplicates with the same information. Note that these commands only identify records that have duplicate values of the sort variables. The records may differ in other respects.

In the acute9x, nacute9x, and dischg9x files, we used:

```
proc sort data=<indata> out=<outdata> nodupkey;
  by scrssn admitday adtime disday sta3n bsinday bsoutday;
```

References

Averill, R. F., McGuire, T. E., Manning, B. E., Fowler, D. A., Horn, S. D., Dickson, P. S., Coye, M. J., Knowlton, D. L., & Bender, J. A. (1992). A study of the relationship between severity of illness and hospital cost in New Jersey hospitals. *Health Serv Res, 27*(5), 587-606.

Barnett, P. G. (1997). Research without billing data. Econometric estimation of patient-specific costs. *Med Care, 35*(6), 553-563.

Barnett, P. G., Chen, S., & Wagner, T. H. (2000). Determining the Cost of VA Care with the Average Cost Method for the 1993-1997 Fiscal Years. *HERC Working Paper #2*.

Calore, K. A., & Iezzoni, L. (1987). Disease staging and PMCs. Can they improve DRGs? *Med Care, 25*(8), 724-737.

Duan, N., Manning, W. G., Morris, C. N., & Newhouse, J. P. (1983). A comparison of alternative models for the demand for medical care. *Journal of Business and Economic Statistics, 1*(2), 115-??

Schneider, D. P., Fries, B. E., Foley, W. J., Desmond, M., & Gormley, W. J. (1988). Case mix for nursing home payment: resource utilization groups, version II. *Health Care Financ Rev, Spec No*, 39-52.

Wagner, T., & Chen, S. (2000). *Economic effects of providing psychosocial residential rehabilitation treatment programs in the VA*. Unpublished manuscript, VA Palo Alto Health Care System.

Appendix A
Reconciliations for 1998

Medicine

692	976 days to Domiciliary

Rehabilitation

452	1 day to medicine
516	2485 days to nursing home
555	171 days to inpatient medicine
603	$219 to inpatient medicine
605	973 days to inpatient medicine
614	51 days to inpatient medicine
642	$2614&122days to inpatient medicine
644	6569 days to nursing home
664	$3968 to inpatient medicine
674	3317 days to blind rehabilitation

Blind Rehabilitation

500	1 day to rehabilitation
537	$109637 to rehabilitation
561	$4737 to rehabilitation
632	221 days to rehabilitation

Spinal Cord

603	$22116 to inpatient medicine
691	$31483 to inpatient medicine

Surgery

620	$346054 to inpatient medicine
655	$240118 to inpatient medicine
687	53 days to inpatient medicine

Inpatient Psychiatry

452	3 days & $1300930 to inpatient med
562	7 days to Intermediate medicine
581	$525 to Intermediate medicine
647	$79619 to inpatient medicine

Substance Abuse

402	828 days to psychiatry
438	312 days to psychiatry
452	$255692 to inpatient medicine
504	251 days to PRRTP
508	916 days to psychiatry
515	691 days to psychiatry
516	7 days to psychiatry
526	23 days to psychiatry
538	6 days to psychiatry
549	1509 days to psychiatry
555	372 days to psychiatry
556	3225 days to psychiatry
558	1 days to psychiatry

Substance Abuse (cont.)

589	$7518 to psychiatry
598	58 days to psychiatry
619	8 days to psychiatry
621	1329 days to psychiatry
622	741 days to psychiatry
631	1681 days to psychiatry
642	$3158493 to psychiatry
646	$1938870 to psychiatry
647	$464271 to inpatient medicine
668	$1635170 to psychiatry
674	67 days to psychiatry
678	1839 days to psychiatry
688	$16686 to psychiatry
691	42 days to psychiatry

Intermediate Medicine

459	$1207930 to nursing home
500	1 day to nursing home
523	6020 days to rehabilitation
529	$110,000 to nursing home
558	211 days to nursing home
589	1207 days to psychiatry
610	8 days to nursing home
612	26673 days to nursing home
642	57 days to nursing home
667	144 days to inpatient medicine
689	$137653 to inpatient medicine

Domiciliary

515	47 days to nursing home
619	$788747 to nursing home
679	420 days to Intermediate medicine

Nursing Home

534	$92298 to inpatient medicine

PRRTP Care

573	6157 days to psychiatry

Reconciliations for 1999

Medicine
- 692 1300 days to domiciliary

Rehabilitation
- 506 20 days to medicine
- 512 7290 days to medicine
- 515 1 days to medicine
- 516 2318 days to nursing home
- 526 321 days to nursing home
- 528 2635 days to nursing home
- 539 2 days to medicine
- 553 275 days to nursing home
- 555 479 days to medicine
- 557 177 days to nursing home
- 605 83 days to nursing home
- 610 1279 days to nursing home
- 623 15 days to medicine
- 642 372 days to medicine
- 644 4925 days to nursing home
- 671 57 days to medicine
- 674 980 days to blind rehab

Blind rehab
- 516 1 days to nursing home

Spinal cord
- 521 3 days to blind rehab
- 670 6 days to medicine

Surgery
- 610 407 days to medicine
- 655 11 days to medicine

Psychiatry
- 529 99 days to medicine

Substance abuse
- 402 575 days to psychiatry
- 438 252 days to psychiatry
- 508 1490 days to psychiatry
- 509 842 days to psychiatry
- 515 286 days to psychiatry
- 516 146 days to psychiatry
- 549 1058 days to psychiatry
- 555 703 days to psychiatry
- 556 3242 days to psychiatry
- 561 279 days to psychiatry
- 589 11 days to psychiatry
- 597 1146492 dollars to psychiatry
- 621 1115 days to psychiatry
- 631 795 days to psychiatry
- 642 3642171 dollars to psychiatry
- 653 2833 days to psychiatry
- 658 12 days to psychiatry
- 664 2 days to psychiatry
- 674 165 days to psychiatry
- 678 1759 days to psychiatry
- 688 342958 dollars to psychiatry
- 691 3 days to psychiatry

Intermediate medicine
- 436 2 days to nursing home
- 504 1 days to nursing home
- 543 106 days to nursing home
- 554 566 days to nursing home
- 558 890 days to nursing home
- 589 1442 days to psychiatry
- 610 1 days to nursing home
- 612 68 days to nursing home
- 619 2001480 dollars to nursing home
- 671 177 days to nursing home
- 678 21 days to nursing home
- 689 3 days to nursing home

Domiciliary
- 515 50 days to nursing home
- 679 352 days to IM

Reconciliations for FY00

Medicine
- 692 moved 1324 days to domiciliary

Rehabilitation
- 512 moved 4925 days to medicine
- 516 moved 1246 days to nursing home
- 526 moved 928 days to nursing home
- 557 moved 146 days to nursing home
- 605 moved 9 days to nursing home
- 608 moved 2452 days to intermediate medicine
- 610 moved 972 days to nursing home
- 612 moved 1 days to nursing home
- 636 moved 1966 days to domiciliary
- 642 moved 1799 days to medicine
- 644 moved 943 days to nursing home
- 668 moved 15 days to nursing home
- 671 moved 50 days to medicine
- 674 moved 23 days to blind rehab

Spinal Cord
- 521 moved 8 days to blind rehab
- 528 moved 21 days to rehabilitation

Surgery
- 610 moved 194 days to medicine
- 655 moved 4 days to medicine

Psychiatry
- 442 moved 1 days to nursing home
- 612 moved 8 days to nursing home

Substance Abuse
- 402 moved 312 days to psychiatry
- 438 moved 140 days to psychiatry
- 508 moved 1026 days to psychiatry
- 509 moved 650 days to psychiatry
- 516 moved 268 days to psychiatry
- 528 moved 92 days to psychiatry
- 531 moved 2812 days to psychiatry
- 549 moved 1 days to psychiatry
- 556 moved 175 days to psychiatry
- 561 moved 433 days to psychiatry
- 585 moved 3219 days to psychiatry
- 621 moved 1216 days to psychiatry
- 630 moved 5828 days to psychiatry
- 631 moved 340 days to psychiatry
- 635 moved 3 days to psychiatry
- 652 moved 397 days to psychiatry
- 653 moved 2397 days to psychiatry
- 656 moved 1 days to psychiatry
- 664 moved 55 days to psychiatry
- 674 moved 69 days to psychiatry
- 678 moved 1865 days to psychiatry

Intermediate medicine
- 508 moved 3 days to nursing home
- 554 moved 1255 days to nursing home
- 558 moved 786 days to nursing home
- 589 moved 2205 days to psychiatry
- 612 moved 77 days to nursing home
- 621 moved 26 days to nursing home
- 671 moved 321 days to nursing home
- 678 moved 60 days to nursing home
- 689 moved 5 days to nursing home

Domicilary
- 512 moved 9016 days to intermediate medicine
- 515 moved 29 days to PRRTP
- 630 moved 17737 days to intermediate medicine
- 673 moved 807 days to nursing home

PRRTP
- 528 moved 23596 days to psychiatry

Medicine
- 459 moved 138 days to NH
- 463 moved 227 days to domiciliary
- 567 moved 93 days to NH
- 679 moved 17 days to IM
- 692 moved 662 days to domiciliary

Rehabilitation
- 512 moved 5011 days to medicine
- 516 moved 2560 days to nursing home
- 526 moved 1141 days to nursing home
- 549 moved 1051 days to nursing home
- 605 moved 1 day to nursing home
- 608 moved 1456 days to intermediate medicine
- 636 moved 4200 days to IM
- 644 moved 27 days to nursing home
- 654 moved 5323 dollars to nursing home
- 664 moved 7646 dollars to nursing home
- 668 moved 5 days to nursing home
- 674 moved 10 days to blind rehab
- 678 moved 20 days to blind rehab
- 679 moved 262 days to NH

Spinal Cord
- 405 moved 10259 dollars to IM
- 644 moved 1 day to NH

Surgery
- 608 moved 54623 dollars to medicine
- 610 moved 421 days to medicine
- 655 moved 1 day to medicine

Psychiatry
- 612 moved 405 days to nursing home

Substance Abuse
- 402 moved 563 days to psychiatry
- 437 moved 393725 dollars to psychiatry
- 438 moved 78 days to psychiatry
- 508 moved 1295 days to psychiatry
- 509 moved 30 days to psychiatry
- 516 moved 170 days to psychiatry
- 542 moved 112313 dollars to psychiatry
- 556 moved 24 days to psychiatry
- 557 moved 276 days to domiciliary (no psych at site)
- 561 moved 366 days to psychiatry
- 570 moved 669 days to psychiatry
- 585 moved 3813 days to psychiatry
- 590 moved 2152 days to psychiatry
- 598 moved 6 days to psychiatry
- 620 moved 5 days to psychiatry
- 621 moved 1030 days to psychiatry
- 630 moved 2506 days to psychiatry
- 632 moved 29 days to psychiatry
- 653 moved 80 days to psychiatry
- 654 moved 155294 dollars to psychiatry
- 657 moved 520 days to psychiatry
- 664 moved 7751 days to psychiatry
- 674 moved 174 days to psychiatry
- 678 moved 1575 days to psychiatry

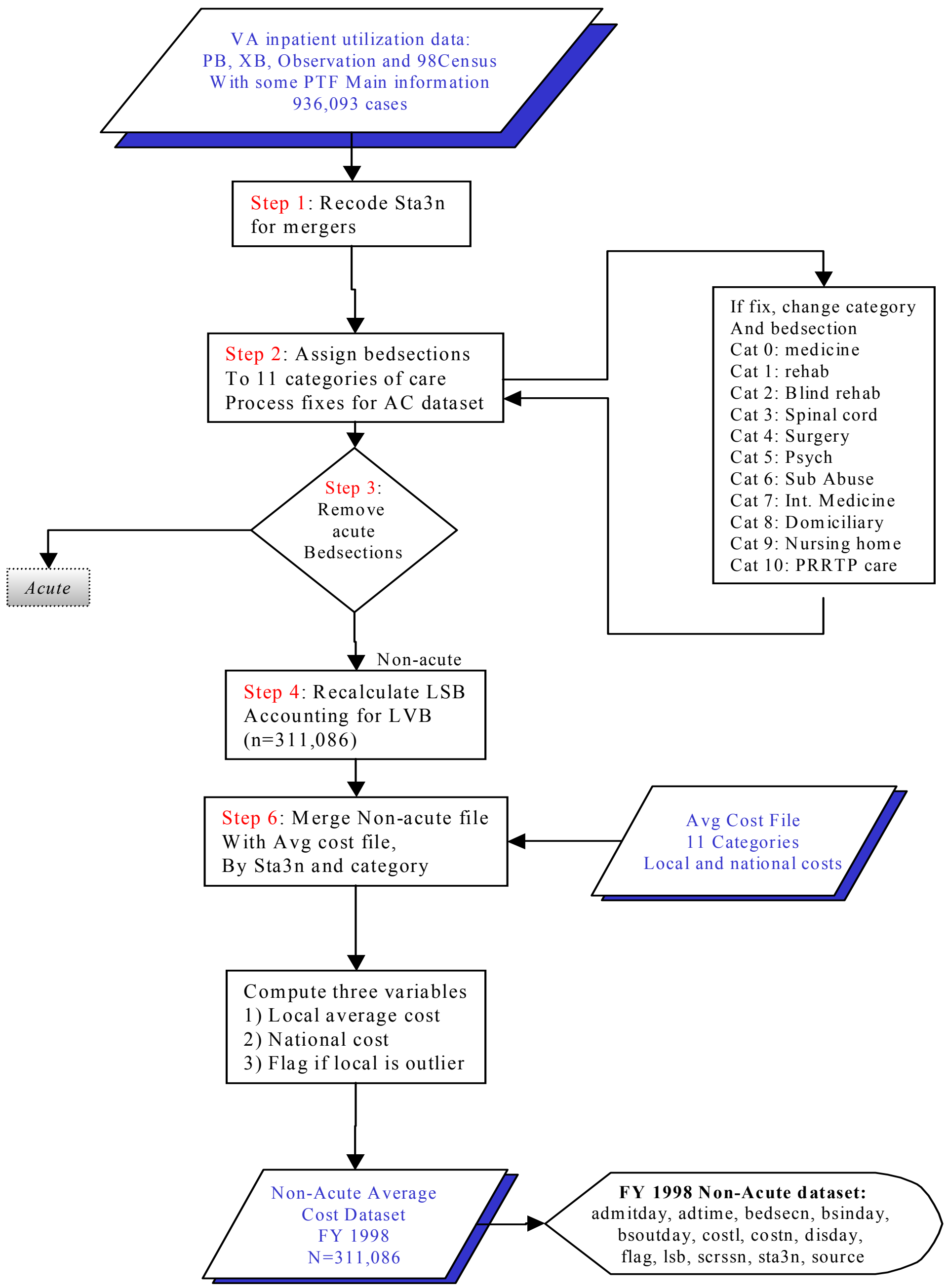
Intermediate medicine
- 554 moved 2997 days to nursing home
- 558 moved 465 days to nursing home
- 656 moved 477 dollars to domiciliary
- 671 moved 237 days to nursing home

Domicilary
- 629 moved 10 days to NH

PRRTP
- 620 moved 48 days to psychiatry

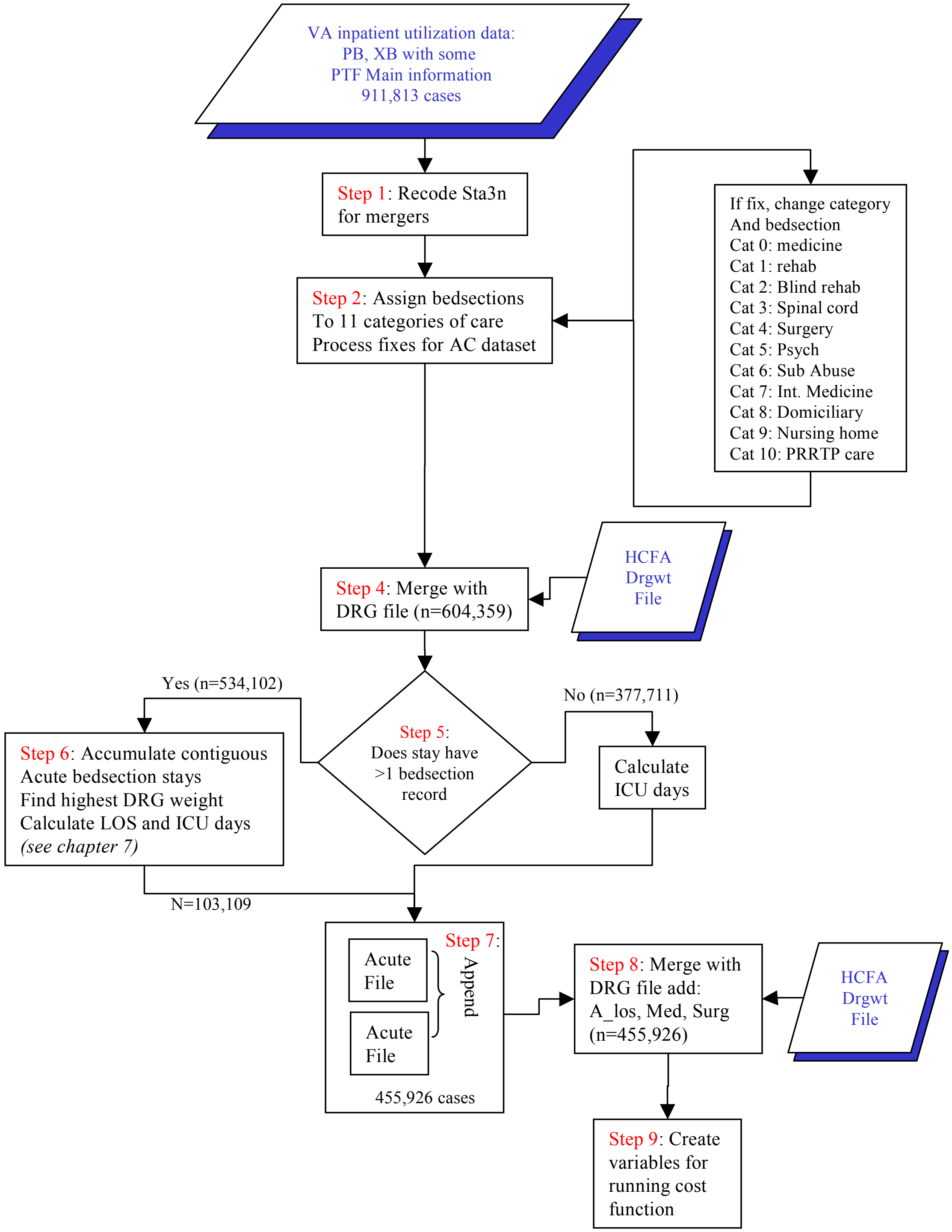
Appendix B
Flow diagram for inpatient care

Development of non-acute average cost dataset



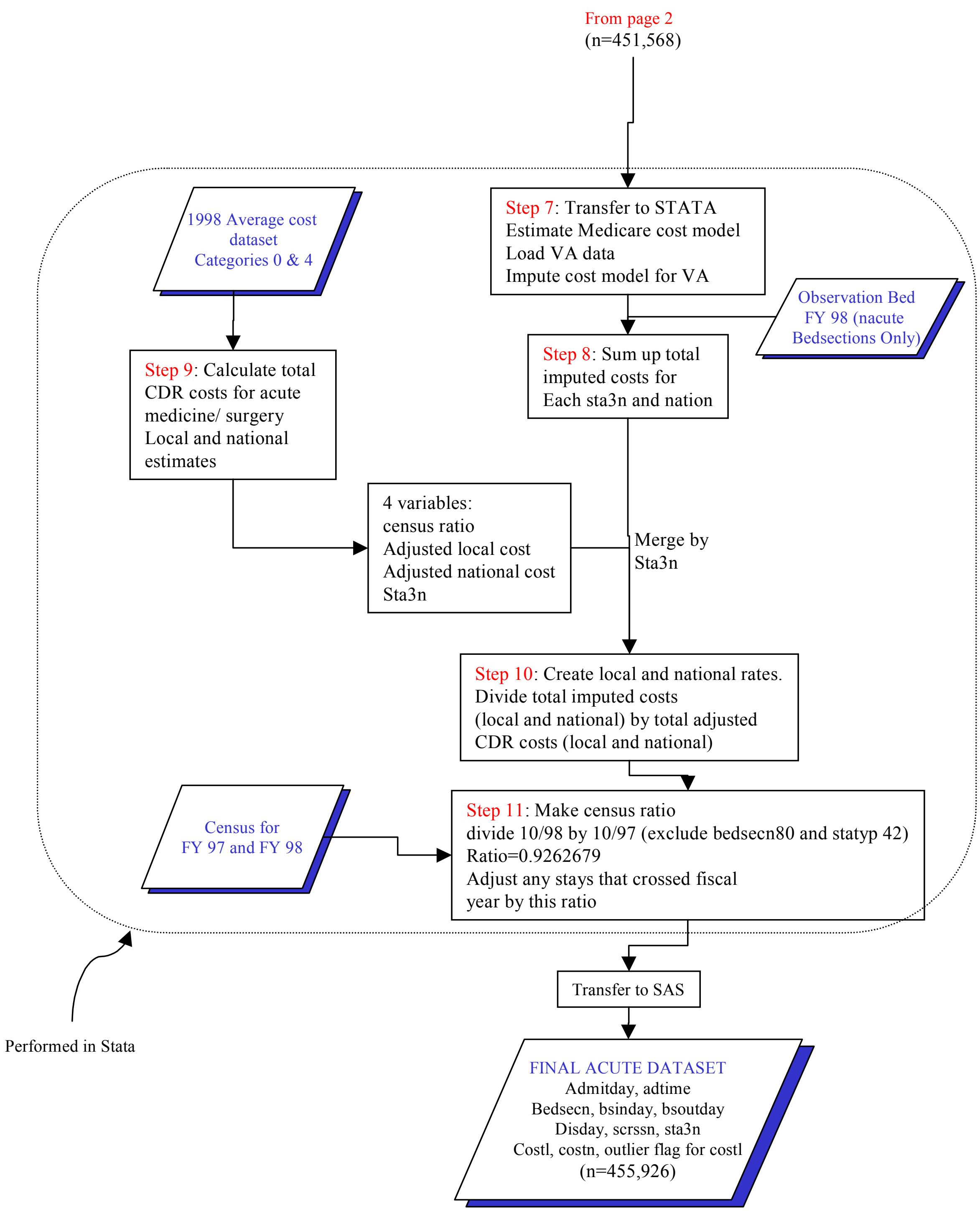
Page: 1

The average cost dataset for acute inpatient costs in FY1998



Page: 2

Scaling the average cost dataset for acute inpatient costs in FY1998



Page: 3

Department of Veterans Affairs **VHA DIRECTIVE 98-025**
Veterans Health Administration
Washington, DC 20420 **May 5, 1998**

RECORDING OBSERVATION AND/OR SHORT-STAY PATIENTS

1. PURPOSE: This Veterans Health Administration (VHA) Directive provides VHA policy for the definition and recording of observation and/or short-stay patients.

2. BACKGROUND: As outlined in the "Vision for Change," VHA will place patients in the most appropriate setting. In many instances, this involves "observing" a patient for an extended period of time without admitting them as an inpatient. While observation units are considered to be outpatient or ambulatory services, current software supporting Nutrition and Food and Pharmacy Services only work for inpatient beds. Properly recording the level of services while maintaining automated support for functional activities will require a creative approach to classifying services to these patients. This policy also complies with current Health Care Financing Association (HCFA) guidelines used in the administration of the Medicare program.

3. DEFINITION

a. Observation Patient. An observation patient is one who presents with a medical condition with a significant degree of instability or disability, and who needs to be monitored, evaluated and assessed for either admission to inpatient status or assignment to care in another setting. An observation patient can occupy a special bed set aside for this purpose or may occupy a bed in any unit of a hospital, i.e., urgent care, medical unit. These types of patients should be evaluated against standard inpatient criteria. These beds are not designed to be a holding area for Emergency Rooms. The length-of-stay in observation beds will not exceed 23 hours.

b. Lodger. A lodger is not an observation patient. By definition a lodger does not receive healthcare services.

NOTE: Routine post-procedure recovery from ambulatory surgery is not observation. Examples: Recovery from a cardiac catheterization and release from the facility within 6 hours of the completion of the catheterization would not constitute post-surgical observation since the normal recovery time is 4 to 6 hours. A patient may report to the medical center for laser removal of cataracts. During the laser procedure, the patient may have a reaction to some of the medication and would be admitted to the appropriate bed section for evaluation of the reaction.

4. POLICY: To accomplish this policy within the context of VHA's supporting software, patients will be assigned to a treating specialty code of Observation. All services and costs associated with Observation treating specialties will be captured and assigned to inpatient services.

THIS VHA DIRECTIVE EXPIRES MAY 5, 2003

5. ACTION

a. The following Patient Treatment File (PTF) Treating Specialties and Cost Distribution

Report (CDR) account numbers are to be utilized for recording Observation patient activity.

Treating Specialty	PTF #	CDR #
Medical Observation	24	1110.00
Surgical Observation	65	1210.00
Psychiatric Observation	94	1310.00
Neurology Observation	18	1111.00
Blind Rehabilitation Observation	36	1115.00
Spinal Cord Injury Observation	23	1116.00
Rehabilitation Medicine Observation	41	1113.00

b. These Treating Specialties should be utilized when setting up Observation Units. The following guidelines and menu options will assist you. Using the Ward Definition menu option create Observation Unit wards. The Treating Specialty should be one of the above Observation Treating Specialties appropriate for the ward location. The service for the Observation Unit ward should be NON-COUNT. Remember to include the Gain and Losses Sheet (G&L) location. Using the Treating Specialty Set-up option, set up the new Treating Specialties.

c. Patients placed on Observation status will be admitted to one of the treating specialties listed above. This will enable the facility to track the patients on the G & L, and use the required Pharmacy and Nutrition and Food Services software to deliver services. An observation patient requiring subsequent admission would be released from Observation status by discharging them from the facility and then admitting them to an acute care-treating specialty.

d. Patients already designated as inpatient status must be discharged and re-admitted to an Observation Treating Specialty for no more than the time limits previously indicated (especially normal ambulatory surgery which are not related to the reason for hospitalization). Following the Observation period, the patient must be re-admitted to inpatient status, if further hospitalization is required. Nursing Home care Unit (NHCU) and Domiciliary (DOM) patients requiring Observation services would be transferred Absent Sick in Hospital (ASIH) from the NHCU or DOM and admitted to an Observation Treating Specialty.

e. Insurance carriers of patients on Observation status will be billed at the appropriate inpatient rate for the medical, surgical or psychiatric bed section using revenue code 760, until such time as an observation unit rate can be established. This is a facility charge and should be billed on an Uniform Billing Form (UB)-92. For billing professional fees only, Current procedural Terminology (CPT) codes should be used. A principal diagnosis should be available for these patients at the time the patient is either discharged and re-admitted to another treating specialty for inpatient care or to an appropriate ambulatory care setting.

f. First party patient charges for Category C observation patients will be billed at the published Category C outpatient visit copayment rate.

g. Utilizing this data report methodology will enable data users to separate the activity of these patients for their purposes. For performance measurement purposes, these patients would NOT be included as acute care inpatients. Procedures performed while a patient is assigned to Observation status will be considered ambulatory for performance measure purposes.

h. Facilities will complete and transmit PTF records for reporting Observation patients when discharged from Observation status. If a patient were admitted following observation, the acute care PTF record would be transmitted after discharge from inpatient care. Attachment A outlines the minimal requirements for patient record documentation of Observation patients.

I. Facilities will complete and transmit PTF records for reporting Observation patients when discharged from Observation status. If a patient were admitted following observation, the acute care PTF record would be transmitted after discharge from inpatient care. Attachment A outlines the minimal requirements for patient record documentation of Observation patients.

j. Patch DG*5.3*176 is being released to implement this directive. Appropriate IB patches will be released in the future.

6. REFERENCES:

Glossary of Healthcare Terms, American Health Information Management Association, 1994, page 14.

7. FOLLOW-UP RESPONSIBILITY

a. For issues affecting classification of patients, Health Administration Service (10C3). Questions concerning classification may be addressed to Kay Evans at (202) 273-8306.

b. For issues concerning billing, Medical Care Cost Recovery (174), Questions concerning billing may be addressed to Nancy Howard at (202) 273-8198.

8. RESCISSIONS: This VHA Directive will expire May 5, 2003.

S/ Thomas Garthwite, M.D.for
Kenneth W. Kizer, M.D., M.P.H.
Under Secretary for Health
Attachment

DISTRIBUTION:	CO:	E-mailed 5/5/98
	FLD:	RD, MA, DO, OC, OCRO and 200 - FAX 5/5/98
	EX:	Boxes 104, 88, 63, 60, 54, 52, 47 and 44 - FAX 5/5/98

Attachment A

Observation Patient Record

Documentation Requirements

Document/ Item	Completion Time	Components of Document Required
Admission Order	On Admission	Timed and dated order for admission of the patient to an Observation Bed
Initial Assessment and History and Physical (H&P)	Immediately	Initial Assessment and screening of physical, psychological (mental) and social status to determine the reason why the patient is being admitted to an Observation Bed, the type of care or treatment to be provided, and the need for further assessment. An extensive Emergency Room (ER) note or Progress Note, documented by the admitting physician, which encompasses the normal criteria for an H&P will suffice as an initial assessment and H&P for the

		Observation patient.
Progress Notes	Within 8 hours - with subsequent notes documented as the patient's condition warrants. 24 hour re-assessments should be documented	Progress Notes should reflect the status of the patient's condition, the course of treatment, the patient's response to treatment and any other significant findings apparent at the time the progress note is documented. Reassessments should include a plan for (1) discharge or transfer; (2) readmission to inpatient status; or (3) continued observation with evaluation and rationale.
Discharge Order	On Discharge	Timed and dated order for discharge from the Observation status.
Discharge Diagnoses	On Discharge	Complete listing of all final diagnoses including complications and comorbidities.
Discharge Note	On Discharge	Summarization of the reason for the Observation admission, the outcome, follow-up plans and patient disposition, and discharge instructions (diet, activity, medications, special instructions). *NOTE: This document may be written in the Progress Notes or dictated, according to local policy.*

Appendix D
Aggregating the bedsection files

The acute98 file was created by aggregating the bedsection file. We wanted details only found in the bedsection file, yet we wanted to create an acute medical-surgical discharge file. The following SAS code was used to aggregate the bedsection file. This code should be used if you want to merge the acute98 file to the bedsection files (PB, XB, & PBO). You can also contact HERC to get an electronic copy of this code.

```
*Step 1: split the file into two files: onebed has all records with only;
*one bedsection stay.  Morebed has all records with more than one stay;


       data onebed morebed;
           set ac4;
           keep acute admitday adtime disday bedsecn bsinday bsoutday drgb drgwt
           lsb ndxb scrssn sta3n age ndxm nsurg sex female died distype ext;
           by scrssn admitday adtime disday sta3n;
           if first.disday and last.disday then output onebed;
           else output morebed;

*sort data;
       proc sort data=morebed out=morebed;
           by scrssn admitday adtime disday sta3n ext;


*Step 2: parse morebed into acute and non-acute stays;
*It then accumulates the acute bedsections when they are contiguous;
       data morebed1;
         set morebed;
         by scrssn admitday adtime disday sta3n ext;


       *Note our definition of acute bedsection is limited by the CDR
       *accounts that we can link to;


       * t_ represents temporary variables to hold information from last case;

       *compute a lag.  This will be used in the accumulation process;
           lagact=lag1(acute);
            if lagact=. then lagact=0;

       retain bedsecn bsoutday bsinday drgb lsb ndxb drgwt icudays t_bsin 0
        t_bsout 0 t_drgb 0 t_lsb 0 t_ndxb 0 t_drgwt 0 t_acute 0 t_icu 0;

       *** t_ represents temporary variables to hold information from last case;


       if not last.disday then do;
             if first.disday then do;
                   if acute=1 then do;
                         if (bedsecn=12 or bedsecn=63) then
                         t_icu=lsb;
                         if (bedsecn~=12 and bedsecn~=63) then
                         t_icu=0;
                         t_bsin=bsinday;
                         t_bsout=bsoutday;
                         t_drgb=drgb;
                         t_lsb=lsb;
                         t_ndxb=ndxb;
                         t_drgwt=drgwt;
                         t_acute=acute;
```

```
        end;
        else if acute=0 then do;
               t_bsin=0;
               t_bsout=0;
               t_drgb=0;
               t_lsb=0;
               t_ndxb=0;
               t_drgwt=0;
               t_acute=0;
               t_icu=0;
        end;
 end;

 else if not first.disday & acute=1 & lagact=0 then do;
        if (bedsecn=12 or bedsecn=63) then t_icu=lsb;
        if (bedsecn~=12 and bedsecn~=63) then t_icu=0;
        t_bsin=bsinday;
        t_bsout=bsoutday;
        t_drgb=drgb;
        t_lsb=lsb;
        t_ndxb=ndxb;
        t_drgwt=drgwt;
        t_acute=acute;
 end;

 else if not first.disday & acute=1 & lagact=1 then do;
        bsinday=t_bsin;
        if t_drgwt>drgwt then drgb=t_drgb;
        if t_drgwt>drgwt then drgwt=t_drgwt;
        if (bedsecn=12 or bedsecn=63) then
        icudays=t_icu+lsb;
        if (bedsecn~=12 and bedsecn~=63) then icudays=t_icu;
        lsb=t_lsb+lsb;
        ndxb=t_ndxb+ndxb;
        *reset values;
        t_drgb=drgb;
        t_drgwt=drgwt;
        t_lsb=lsb;
        t_ndxb=ndxb;
        t_bsout=bsoutday;
        t_icu=icudays;
 end;

 else if not first.disday & acute=0 & lagact=0 then do;
        t_bsin=0;
        t_bsout=0;
        t_drgb=0;
        t_lsb=0;
        t_ndxb=0;
        t_drgwt=0;
        t_acute=0;
        t_icu=0;
 end;

 else if not first.disday & acute=0 & lagact=1 then do;
        bsinday=t_bsin;
        bsoutday=t_bsout;
        drgb=t_drgb;
        lsb=t_lsb;
        ndxb=t_ndxb;
        drgwt=t_drgwt;
        acute=t_acute;
        icudays=t_icu;
        t_bsin=0;
        t_bsout=0;
        t_drgb=0;
        t_lsb=0;
```

```
                t_ndxb=0;
                t_drgwt=0;
                t_acute=0;
                t_icu=0;
                output;
            end;
        end;
        else if last.disday then do;
            if (acute=1 & lagact=1) then do;
                if (bedsecn=12 or bedsecn=63) then
                icudays=t_icu+lsb;
                if (bedsecn~=12 and bedsecn~=63) then icudays=t_icu;
                bsinday=t_bsin;
                bsoutday=bsoutday;
                if t_drgwt>drgwt then drgb=t_drgb;
                drgwt=max(t_drgwt,drgwt);
                lsb=t_lsb+lsb;
                ndxb=t_ndxb+ndxb;
                output;
            end;

            if (acute=1 & lagact=0) then do;
                if (bedsecn=12 or bedsecn=63) then icudays=lsb;
                if (bedsecn~=12 and bedsecn~=63) then icudays=0;
                bsinday=bsinday;
                bsoutday=bsoutday;
                drgb=drgb;
                lsb=lsb;
                ndxb=ndxb;
                drgwt=drgwt;
                acute=acute;
                output;
            end;

            else if (acute=0 & lagact =1) then do;
                bsinday=t_bsin;
                bsoutday=t_bsout;
                drgb=t_drgb;
                lsb=t_lsb;
                ndxb=t_ndxb;
                drgwt=t_drgwt;
                acute=t_acute;
                icudays=t_icu;
                output;
            end;
            else do;
                t_bsin=0;
                t_bsout=0;
                t_drgb=0;
                t_lsb=0;
                t_ndxb=0;
                t_drgwt=0;
                t_acute=0;
                t_icu=0;
            end;
        end;


         data morebed2;
             set morebed1;


*Step 3: compile the accumulated acute stays with the onebed acute stays;
         data onebed1;
           set onebed;
           if acute=1;
           if (bedsecn=12 or bedsecn=63) then icudays=lsb;
```

```
   if (bedsecn~=12 and bedsecn~=63) then icudays=0;

  *the complete acute file;
data allbed;
          set morebed2 onebed1;
          keep acute admitday adtime disday bsoutday drgb
       drgwt lsb scrssn sta3n icudays ndxm female died
       age distype;

data allbed (replace=yes);
          set allbed;
          *remove discharges in prior year (BSOUTDAY HAS BEEN ALTERED);
          if bsoutday>mdy(9,30,97);
```